CUK-H05960-29-P7207

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

Education

FEBRUARY 1990

29

VOL. 60, PART 1

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by

Professor CHARLES DESFORGES

Assistant Editors:	Sub-Editor:
Dr H. C. M. CARROLL	Dr S. CUNNINGHAM
Dr D. FONTANA	
Mr G. HARVARD	Book Review Editor:
Dr P. TOMLINSON	FLORA J. MACLEOD
Dr R. YOUNGMAN	

Advised by the Editorial Committee:

Professor G. BROWN
Ms. I. LUNT
Dr A. MANSTEAD
Professor P. W. ROBINSON
Dr K. WHELDALL

P, 7207

Vol. 60, Part 1

The *British Journal of Educational Psychology* is issued by the British Psychological Society once a term, in February, June and November. The subscription is £30.00 a year ($60.00 North America), or £10 50 each number.

Subscriptions and Orders should be sent to Scottish Academic Press, 139 Leith Walk, Edinburgh, EH6 8NS. Members of the British Psychological Society receive the *Journal* on special terms.

Papers for publication in duplicate should be sent to Professor Charles Desforges, University of Exeter, St Luke's, Heavitree Road, Exeter, EX1 2LU. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for The *British Journal of Educational Psychology* should be sent to. Flora Macleod, University of Exeter, St Luke's, Heavitree Road, Exeter, EX1 2LU.

Advertisements for this *Journal* should be sent to Scottish Academic Press.

Published by
SCOTTISH ACADEMIC PRESS, 139 Leith Walk, Edinburgh, EH6 8NS,
for the British Journal of Educational Psychology, Ltd.

All rights reserved

© 1990. The British Journal of Educational Psychology

PRINTED IN GREAT BRITAIN BY
BELL AND BAIN LTD., GLASGOW

Br. J. educ. Psychol., 60, 1-9, 1990

-1505
LB777
370·1505
B777

INDIVIDUALISED INSTRUCTION, CONGRUENT AUTHORITY STRUCTURE AND STUDENT OUTCOMES

BY YOHANAN ESHEL AND JENNY KURMAN
(University of Haifa, Israel)

SUMMARY. Measures of formal and informal authority structure were employed in order to assess academic and non-academic outcomes of individualised instruction. A sample of 386 fourth and fifth-grade elementary school children studying in nine individualised and 10 traditional classes was employed. Ths academic outcomes of either method of instruction were found to be contingent on the congruence between formal and informal authority structure: higher achievement scores were obtained when individualistion as school policy was accompanied by teacher sharing of authority with pupils at the informal classroom level. Achievement in less individualised classes was higher when teacher did not extend decision-making opportunities to pupils. Non-academic pupil outcomes were generally unrelated to either formal or informal authority structure. Similarities and differences between these results and former findings are discussed in light of cross-cultural differences in classroom individualisation.

INTRODUCTION

INDIVIDUALISED instruction is one of several methods used to present an alternative to traditional methods of education that is more child centred. A conceptualisation of individualised instruction developed by Howes (1970) divided it into three major modes of teaching and learning. Two of these are of relevance to the present discussion. In one mode, termed "differentiated instruction", "emphasis is placed on accounting for individual differences in designating teaching and carrying on its purpose. Such instruction has as its core not only a focus on the variety and range of individual differences but it also emphasises the role of the teacher in providing for, dealing with and meeting or accommodating instruction to these differences" (pp. 72-73). Howes used the term "'individualisation" for a second mode of instruction "in which a person can investigate and discover for himself; a mode of teaching which increases learner direction and control; and management procedures and arrangements which match individual development with larger and larger amounts of self-selection, self-guidance and learner responsibility" (pp. 76-77).

We believe that the importance of Howes' analysis lies in raising the question, is there a one-to-one relationship between the application of "differentiated instruction" and attaining outcomes promoted by "individualisation". In other words, is it conceivable that instruction will be congruent with pupil needs without raising their academic achievement, independence, or responsibility?

Several reviews of the available research examine this issue in an attempt to present a comprehensive evaluation of the effectiveness of more open educational methods. These studies fail to reach an unequivocal conclusion (Horwitz, 1979; Peterson, 1979; Giaconia and Hedges, 1982). On the whole, researchers tend to agree that more open educational methods are somewhat more effective than traditional instruction for non-achievement outcomes. Traditional education is somewhat more effective in promoting academic achievement. Similar results of no consistent advantage of individualised instruction were also reported by studies employing a more narrow definition of this method (Royce and Shank, 1975; Miller, 1976; Schoen, 1976; Hartley, 1977; Bangert *et al.*, 1983; Slavin *et al.*, 1984).

The inconsistent empirical results in this domain have been attributed to two major factors: (a) vast differences in operationalisation of open education and (b) ambiguous conceptualisation of classroom openness (Marshall, 1981). Evidence on implementation suggests that teacher attitudes and beliefs, societal norms and student body characteristics often lead individualised, open and progressive classes to retain a variety of traditional elements as integral parts of their instructional method (Berlak *et al.*, 1975; Walberg and Thomas, 1975; Elliott, 1977; Klein and Eshel, 1980; Harrison and Glaubman, 1982; Eshel and Klein, 1984).

This diversity in implementation is often accompanied by ambiguous conceptualisation (Hill, 1975) and abundance of operational definitions of open methods (Traub *et al.*, 1972; Walberg and Thomas, 1972). Hence, no agreement has been established among students of the field as to which are the more pertinent elements of individualised instruction and which are of lesser importance: richness of learning opportunities and graded learning materials, recognising students' right to choose their own academic course, turning the teacher into a facilitator of learning rather than its conductor, development of independent thinking, student responsibility and autonomy, or flexibility of the physical and social environment of the classroom. It is quite clear, therefore, that a better understanding of individualised instruction and its relationship to student outcomes requires a more crisp definition of the major components of the intervention involved.

In line with the Epstein and McPartland (1979) model it will be assumed that formal as well as informal authority structures are essential components of classroom individualisation. The fact that these structures have been practically ignored by the vast majority of studies on individualisation may account for some of the inconsistent results in this field. According to the reasoning implied by these structures, the ultimate aim of attempts to individualise education is to change the traditional division of power between teachers and students at the formal as well as the informal levels of authority in the class. Formal authority structure relates to the organisation of instruction as determined by the educational policy of the school. The informal aspect of authority structure reflects the extent to which the decision making prerogative is either shared with the students or reserved as the sole privilege of the teacher (for instance, are children allowed to determine whether to study individually or in groups, or if they may choose the group they prefer to work with?).

Epstein and McPartland (1979) demonstrated that the correlation between formal and informal aspects of authority structure is rather small so that both teachers who share authority with students and others who do not can be found in every school. Studies of authority structure have found that shared authority is associated with higher student self-confidence, satisfaction and adjustment to school, and with more positive attitudes toward school, peers and teachers. In some cases it was also correlated with academic achievement (Wang and Stiles, 1976; Epstein and McPartland, 1979; Ritcher and Tjosvold, 1980; Epstein, 1981).

The present study investigates outcomes of individualised instruction in Israel. This version of the individualised classroom devised, conducted and supervised by the Centre for Educational Technology, aims at attaining "differential instruction" as well as "individualisation": pupils are provided with a variety of graded learning materials and learning activities. They are encouraged to assume responsibility and determine their own rate of learning as well as the choice of learning assignments, provided that they master a basic core curriculum. Learning is carried out individually or in small flexible groups with occasional whole class teaching, allowing pupils greater opportunity to develop according to their inclinations and learning styles. The philosophy of individualised instruction maintains that these conditions will increase teachers' belief in pupils' potential for growth as well as their job satisfaction, and lead pupils to improved performance and greater realisation of potentials and abilities. The present study examines the more specific assumption that effectiveness of this method is a joint function of its formal and informal authority structure: more positive educational outcomes are contingent on the congruity between these two levels of authority structure.

Agreement between formal and informal messages conveyed by school and teacher provides students of individualised instruction with both the means and the legitimisation to become more responsible for their learning and to choose the course of their academic development. Lack of congruity between formal and informal messages of individualisation may confuse and frustrate students and lead them to suspect teachers' reliability and sincerity. It may, therefore, create an atmosphere detrimental to attaining both academic and non-academic aims of individualised instruction.

The same reasoning seems to apply to traditional instruction. Congruence between formal and informal authority structure in these schools clearly specifies roles for teacher and for pupil, while attempts of individual teachers to change these role definitions which are not accompanied by a change in school policy are liable to be confusing and counter-productive. It will be hypothesised, therefore, that congruent authority structure in either open or traditional classes is associated with the following student outcomes: (a) higher academic attainment, (b) more positive self-esteem and (c) lower school anxiety, dependence and social desirability.

It will be further hypothesised that the two groups of schools also differ on several additional variables. Compared to their more traditional peers, teachers in individualised schools will (a) extend more decision-making opportunities to pupils, (b) have more confidence in pupils' ability and motivation, (c) have a more liberal control ideology, (d) have more autonomy in their schools, (e) describe their schools' climate in more positive terms, and (f) enjoy higher job satisfaction.

METHOD

Sample

The sample consisted of 386 school pupils who studied in all 19 fourth and fifth grade classrooms in seven schools. Three of the schools had achieved high competence in employing individualised instruction, while the four others were traditional schools. All seven schools were designated as disadvantaged by the Israeli Ministry of Education. This designation means that the vast majority of pupils are low achievers who come from lower-class families of Afro-Asian origin. A disadvantage index was computed for each pupil by summing up the following attributes: parents with less than high school education, parents of Afro-Asian (rather than European-American) origin and four or more children in family.

The extent of school individualisation was determined by experts of the Centre for Educational Technology, who developed and implemented this method throughout Israel. Two experts were employed who had gained intimate knowledge of the schools and their progress from periodic consultation visits to the schools over the course of several years. All the disadvantaged schools in two districts in Israel that had adopted the individualised instruction method were rank ordered according to their level of individualisation. Each school was rated on a three-point scale (1 = low, 2 = medium and 3 = high) by the two experts with regard to 11 dimensions. Extent of school individualisation was determined by summing up these ratings. The dimensions employed represented the major elements of the Israeli version of individualised instruction: (1) designation of both mutual and individual learning assignments; (2) pupils learn at their own rate; (3) diversified teaching aids available; (4) diversified pupil activities available; (5) various teaching and learning procedures employed; (6) diversity and flexibility of learning groups; (7) texts and materials are graded; (8) pupils choose learning assignments; (9) independent planning of daily agenda by pupils; (10) pupils assess and monitor their own learning; (11) systematic assessment of pupils' progress carried out. It is important to note that ratings of all these categories represent overall estimation of the educational beliefs and instructional policies of the school, rather than individualisation scores assigned to specific classrooms in it. They represent, therefore, the formal authority structure of the investigated schools.

The present sample included three of the four schools that scored highest on these dimensions and four traditional schools of similar ethnic and SES composition.

Instruments

I Pupil academic outcomes

Arithmetic, reading comprehension and vocabulary tests were employed. Arithmetic attainment was assessed by the Centre for Eductional Technology tests for grades 3 and 4. These tests were employed since a pretest had indicated that pupils in the sample lagged way behind the required curriculum and many of them were unable to answer arithmetic items devised for their grade level. Each test consisted of 50 items representing the curriculum required by the Ministry of Education. The split-half reliabilities of these tests were r = 0.88 for the 3rd grade and r = 0.87 for the 4th grade test.

Reading comprehension was examined by 28 items that had been included in a nationwide study evaluating primary education (Minkovich *et al.*, 1980). The split-half reliability of this test was r = 0.723.

The vocabulary test was derived from the MILTA learning ability test (Ortar, 1966). The instrument consisted of two parallel forms of 25 items each. A former study indicated that the correlation between this measure and a reading comprehension test was r = 0.64 (Eshel *et al.*, 1987). The correlation between the two forms of this test was r = 0.96. The vocabulary test correlated r = 0.91 with the entire MILTA test.

In the following analyses raw scores of all the three academic measures are transformed into standardised scores computed separately for each grade level, with a mean of 0 and a standard deviation of 1.

II Pupil non-academic outcomes

Three personality scales were employed: school anxiety, a 10-item scale including items such as "I'm afraid of exams," or "I'm totally confused when teacher asks me questions unexpectedly" (Cohen, 1976); social desirability, measured by 12 items derived from Crandall *et al.* (1965) scale ("I'm always respectful of older people", "I never forget to say 'please' and 'thank you'"); dependency, estimated by eight items from the Student Learning Style Scale (Reichman and Grasha, 1974) ("I try my best to write in my notes everything teacher says", "Before working on a class project I try to get the approval of the instructor"). The internal consistency of the non-academic scales was assessed using Cronbach's alpha index. The alpha values for these three scales were 0.619, 0.666 and 0.678 respectively.

Academic self-concept, A seven-item self-description questionnaire adapted from the Minkovich *et al.* (1980) study ("I obtain high scores in class", "I know the right answers to questions asked by teacher") (alpha = 0.858).

III Teacher measures

Pupil control ideology. A 20-item instrument (alpha = 0.662), devised by Willower *et al.* (1967) ("Pupils should not be allowed to contradict teacher in class").

Extended authority scale. A 12-item scale determining the extent to which teachers are willing to share decision-making with their pupils ("Pupils determine classroom rules and behaviour", "Pupils voluntarily group and regroup themselves", "Pupils set their own pace in class-work") (Eshel *et al.*, 1982; alpha = 0.759).

Job satisfaction. A 10-item instrument estimating teachers' satisfaction with their professional development, relations with students or student achievement (alpha = 0.694).

School climate. Three sub-scales derived from the School Organisational Scale (Zak, 1981) were employed to measure teacher autonomy ("I feel free to act according to my educational convictions in class", four items, alpha = 0.613), peer relations ("There is a lot of tension among teachers in my school", five items, alpha = 0.516), and general school climate ("Teachers in my school are highly motivated", four items, alpha = 0.683).

Pupil assessment. Teachers rated each of their pupils on two dimensions using 5-point scales, (a) motivation to learn, and (b) ability and understanding. Each pupil was also graded for his or her achievements in arithmetic and reading comprehension on rating scales ranging from 40 to 100.

Procedure

Teachers completed their questionnaires individually. Pupil measures were group administered on the last month of the school year, to all pupils present in their classes on testing day.

RESULTS

Teacher outcomes

Schools were assigned to the individualised or the traditional category on the basis of expert ratings. Teacher questionnaires were employed in order to validate these ratings and gain further understanding of the differences between the two methods of instruction from the teacher's perspective. Teachers' feelings, beliefs and behaviour in individualised classes (N = 9) and in more conventional ones (N = 10) were compared using t tests. Table 1 presenting these data submits that individualised-instruction teachers are more satisfied with their work (t = 2.73, P <0.017), feel they have more autonomy in planning their teaching (t = 2.29, P <0.037) and are willing to grant their pupils greater control over their own learning (t = 3.30, P <0.004). In keeping with this point of view they are also inclined to believe that their pupils perform higher than the national average level while more traditional teachers believe their pupils to score lower than this average (t = 2.37, P <0.033).

TABLE 1

MEANS, STANDARD DEVIATIONS AND T-TESTS COMPARING TEACHER PERSPECTIVES ON INDIVIDUALISED INSTRUCTION AND TRADITIONAL TEACHING

	Method of Instruction				
	Traditional		Individualised		
Variable	M	SD	M	SD	t
Pupil decision making	2.01	(0.35)	2.52	(0.34)	3.30**
Teacher autonomy	3.97	(0.47)	4.36	(0.31)	2.23*
Job satisfaction	3.66	(0.24)	4.05	(0.36)	2.73*
Estimated student ability	2.44	(0.53)	3.13	(0.64)	2.37*
School atmosphere	3.91	(0.49)	4.14	(0.22)	NS
Peer relations	3.66	(0.59)	3.89	(0.40)	NS
Control ideology	2.89	(0.32)	2.71	(0.41)	NS

*P < 0.05; **P < 0.01

CENTRAL LIBRARY

No significant differences were found for other variables that were supposed to distinguish between the two groups of teachers: pupil control ideology and three aspects of school climate (teacher-principal relations, peer relations, and general school atmosphere).

Pupil outcomes

The effects of individualised instruction and of shared authority on pupil outcomes were examined separately for academic and non-academic variables, by means of multiple analyses of variance (MANOVA), with age and sex as the constant factors, and with teacher effects nested within method of instruction. The MANOVA results indicated significant overall effects of both method of instruction (F (3, 12) = 4.22, P <0.03; Wilks' L = 0.486), and the interaction between method of instruction and extended authority (F (3, 364) = 2.97, P <0.04; Wilks' L = 0.976) on academic achievement. No significant overall main effect was found for the extended authority scores.

Examination of mean academic achievement scores indicated that individualised instruction was associated with higher performance on reading comprehension and on the vocabulary tests. Table 2 presenting the MANOVA data maintains that both these differences are significant (F (1, 14) = 6.43, P <0.02, and F (1, 14) = 5.44, P <0.04 accordingly).

TABLE 2

F VALUES OF MANOVA ANALYSES OF THE EFFECTS OF METHOD OF INSTRUCTION AND EXTENDED AUTHORITY SCORES ON STUDENT OUTCOMES, WITH GRADE LEVEL AND SEX AS THE CONSTANT VARIABLES AND TEACHER EFFECTS NESTED WITHIN METHOD OF INSTRUCTION

Academic	Method of instruction		Extended authority		Method x authority	
Variables	MS	F(1.14)	MS	F(1.377)	MS	F(1.377)
Arithmetic	0.27	NS	0.34	NS	6.34	7.37**
Reading comprehension	8.27	6.43*	0.04	NS	1.67	NS
Vocabulary	14.86	5.44*	2.21	NS	4.54	5.28*
Non-academic Variables	MS	F(1.14)	MS	F(1.385)	MS	F(1.385)
Academic self-concept	0.40	NS	2.73	5.88*	0.03	NS
School anxiety	0.24	4.98*	0.07	NS	0.02	NS
Social desirability	0.12	NS	0.03	NS	0.06	NS
Dependence	0.07	NS	0.04	NS	0.11	NS

*P < 0.05; **P < 0.01

Further examination of the means suggested that in individualised classes, greater pupil responsibility for learning was associated with higher academic achievement. Similar attempts to increase the pupils' share in decision-making in more traditional classes were either unrelated or negatively related with these achievement scores.

The MANOVA data indicated that these interaction effects were significant for the arithmetic test F (1, 385) = 7.37, P <0.01, and for the vocabulary score F (1,385) = 5.28, P <0.05. It seems that a high level of individualised instruction at the formal level of authority structure would lead more often to higher academic achievement when decision-making is not the sole prerogative of the teacher. It is further suggested that granting responsibility to traditional-school pupils may be of no use to them if they are unwilling or unable to use it competently.

Individualised instruction was assumed to promote several non-academic pupil outcomes. The MANOVA data present little support for this contention. Only method of instruction had significant overall effect on these variables F (4, 11) = 4.23, $P < 0.03$, Wilks' L = 0.394). Table 2 shows that this effect is accounted for by differences in school anxiety between methods of instruction F (1,14) = 4.98, $P < 0.05$. Examination of the means indicated lower anxiety in the less individualised classes.

Pupils' achievement, ability and motivation were also estimated by teachers. Individualised instruction teachers tended to reveal more faith in their pupils than traditional peers, rating them as having higher academic ability F (1, 14) = 5.01, $P < 0.05$ and higher motivation F (1, 14) = 9.33, $P < 0.01$. This faith was not evident in their ratings of pupils' arithmetic and language achievement, perhaps because pupils were evaluated relative to other classmates. None of the teacher assessments was found to be significantly associated with shared authority.

DISCUSSION

Congruent authority structure

The present study examines the relationship between pupil outcomes and measures of individualised instruction, employing an analysis of formal and informal authority structure. The former relates to methods of instruction determined by school policy. The latter reflects the extent to which the teacher is willing to grant pupils a meaningful degree of independence in the decision-making process. In line with Epstein and McPartland (1979) theorising it was assumed that student outcomes are a joint function of these two definitions of individualisation.

The MANOVA results presented above indicated that this indeed is the case as far as academic achievement is concerned. Significant interaction effects were found for both arithmetic and vocabulary scores. Achievement was consistently higher when there was congruence between the formal and the informal levels of individualisation. In the more individualised schools both arithmetic and vocabulary test scores were higher when teachers extended decision-making rights to the pupils. Achievement in these schools was considerably lower when the teacher retained decision-making as his or her exclusive prerogative. In less individualised schools the opposite seems to hold. An attempt to share authority with pupils that was not accompanied by a parallel change of the formal authority structure was associated with small improvement in arithmetic and vocabulary performance, and with lower reading comprehension scores. It may be concluded then, that although individualised instruction may serve as a vehicle for improved learning, it does not consist of a sufficient condition for attaining this end. Its effectiveness seems to depend on the extent to which it conveys an unambiguous message delineating the role of students.

Incongruence between the formal and informal levels of individualisation may be perceived by pupils as an insincere double message: requesting greater responsibility without increasing their own authority over their learning. The data show that their academic potential is, in fact, more fully realised when they have reason to believe that the messages conveyed in class are reliable and sincere statements, not deviating from declared school policy.

Non-academic pupil outcomes

The present findings of insubstantial relations between individualised instruction and non-academic outcomes are consistent with former Israeli studies (Levin *et al.*, 1981). However, they agree neither with American studies investigating alternative education in general (Horwitz, 1979; Peterson, 1979; Giaconia and Hedges, 1982) nor with more specific research into formal and informal structure of authority (Wang and Stiles, 1976; Epstein and McPartland, 1979; Ritcher and Tjosvold, 1980; Epstein, 1981).

There is reason to believe that these differences in research outcomes reflect the specific characteristics of the educational system in Israel. This system has been described (Kleinberger, 1969) as being quite traditional in the emphasis on the acquisition of content, being elitist in its demand for high level performance, and expecting collective identification with national goals and purposes.

It is quite evident that this educational perspective is not in accordance with the declared aims of individualised instruction. Teachers in innovative schools face therefore an almost inevitable conflict between these two dissonant value systems (Elliott, 1977). There is reason to believe that inconsistent non-academic student outcomes may reflect, in the present case, the variance of teacher solutions for this role conflict.

Educational outcomes reflect quite often differences between teachers rather than differences between methods of instruction. We believe, however, that the nested design employed in this study controlled reasonably well for teacher effects, so that the present data reflect primarily effects of formal and informal authority structure. These results, if this indeed is the case, are a step toward a better understanding of the relationships between authority structure and both intellectual and emotional outcomes of schooling.

Correspondence and requests for reprints should be addressed to Dr Yohanan Eshel, Department of Psychology, University of Haifa, Mount Carmel, Haifa 31999, Israel.

REFERENCES

Bangert, R. L., Kulik, J. A., and Kulik, C. C. (1983). Individualized systems of instruction in secondary schools. *Rev. educ. Res.*, 53, 143-158.

Berlak, A., Berlak, A., Bagenstos, N., and Mikel, F. (1975). Teaching and learning in English primary schools. *School Rev.*, 83, 215-243.

Cohen, L., (1976). *Educational Research in Classrooms and Schools*, London: Harper and Row.

Crandall, V. C., Crandall, V. S., and Katkovsky, W. A. (1965). Children's social desirability questionnaire. *J. consult Psychol.*, 29, 27-36.

Elliott, J. (1977). Developing hypotheses about classrooms from teachers' practical constructs: an account of the work of the Ford Teaching Project. *Interchange*, 7, 2-22.

Epstein, J. L. (1981). Patterns of classroom participation, student attitudes and achievements. In Epstein, J. L. (Ed.), *The Quality of School Life*. Lexington, Mass.: Heath.

Epstein, J. L., and McPartland, J. M. (1979). Authority structures. In Walberg, J. J. (Ed.), *Educational Environments and Effects: Evaluation, Policy and Productivity*. Berkeley, Calif.: McCutchon.

Eshel, Y., and Klein, Z. (1984). School desegregation and achievement, In Amir, Y. and Sharan, S. (Eds.), *School Desegregation, Cross Cultural Perspective*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Eshel, Y., Shachar, H., and Wainryb, C. (1982). *Pupils' Perspectives of Open Education*. Research Report submitted to the Ministry of Education, Jerusalem.

Eshel, Y., Wainryb, C., and Shachar, H. (1987). Open education and reading comprehension of high-ability and low-ability elementary school pupils. *Br. J. educ. Psychol.*, 57, 166-178.

Giaconia, R. M., and Hedges, L. V. (1982). Identifying features of effective open education. *Rev. educ. Res* , 52, 579-602.

Harrison, J., and Glaubman, R. (1982). Open education in three societies. *Compar. Educ. Rev.*, 76, 352-373.

Hartley, S. S. (1977). Meta-analysis of the effects of individually paced instruction in mathematics. (Doctoral Dissertation, University of Colorado, 1977). *Dissert. Abstr. Int.*, 38, 4003.

Hill, B. V. (1975). What's "open" about open education? In Nyberg, D. (Ed.), *The Philosophy of Open Education*. London: Routledge and Kegan Paul.

Howes, V. M. (1970). Individualized instruction: form and structure. In Howes, V. M. (Ed.), *Individualization of Instruction*. London: Macmillan.

Horwitz, R. A. (1979). Psychological effects of the "open classroom". *Rev. educ. Res.*, 49, 71-86.

Klein, Z. and Eshel, Y. (1980). *Integrating Jerusalem Schools*. New York: Academic Press.

Kleinberger, A. F. (1969). *Society, Schools and Progress in Israel*. Oxford: Pergamon.

Levin, T., Rubinek, B., and Amiad, R. (1981). Student perception of classroom climate. *Stud in Educ.*, 29, 83-96.

Marshall, H. H. (1981). Open classrooms: has the term outlived its usefulness? *Rev. educ. Res.*, 51, 181-192.

Miller, R. L. (1976). Individualized instruction in mathematics: a review of research. *Math. Teacher*, 345-351.

Minkovich, A., Davis, D., and Bashi, Y. (1980). *An Evaluation Study of Israeli Elementary Schools*. Jerusalem: Magnes Press.

Morrison, T. (1974). Control as an aspect of group leadership in classrooms. *J. educ. Psychol.*, 56, 38-64.

Ortar, G. (1966). *The MILTA Test*. Jerusalem: School of Education, Hebrew University.

Peterson, P. L. (1979). Direct instruction reconsidered. In Peterson, P. L., and Walberg, H. J. (Eds.), *Research on Teaching: Concepts, Findings and Implications*. Berkeley, Calif.: McCutchon.

Reichman, S., and Grasha, A. (1974). A rational approach to developing and assessing the construct validity of Student Learning Style Scale instrument. *J. Psychol.*, 87, 213-223.

Ritcher, F. D., and Tjosvold, D. (1980). Effects of student participation in classroom decision making on attitudes, peer interaction, motivation and learning. *J. appl. Psychol.*, 65, 74-80.

Royce, G. L., and Shank, J. (1975). Scorecard for individualized instruction. *Science Teacher*, 42, 27-29.

Schoen, H. L. (1976). Self-paced mathematics instruction: how effective has it been? *Arith. Teacher*, 90-96.

Slavin, R. E., Madden, N. A., and Leavey, M. (1984). Effects of comparative learning and individualized instruction on mainstream students. *Except. Children*, 50, 434-443.

Traub, R. E., Weiss, J., Fisher, C. W., and Musella, D. (1972). Closure on openness: describing and quantifying open education. *Interchange*, 3, 69-83.

Walberg, H. J., and Thomas, S. C. (1972). Open education: an operational definition and validation in Great Britain and United States. *Am. educ. Res. J.*, 9, 197-208.

Walberg, H. J., and Thomas, S. C. (1975). An analysis of American and British open education. In Spodek, B. and Walberg, H. J. (Eds.), *Studies in Open Education*. New York: Agathon Press.

Wang, M. C., and Stiles, B. (1976). An investigation of children's concept of self-responsibility for school learning. *Am. educ. Res. J.*, 13, 159-179.

Willower, D., Eidell, T., and Hoy, W. (1967). *The School and Pupil Control Ideology*. Pennsylvania State University Press.

Zak, I. (1981). The evaluation of school organization climate. In Levy, A. and Nevo, D. (Eds.), *Evaluative Roles in Education*. New York: Gordon and Breach.

(*Manuscript received* 30*th January*, 1989)

Br. J. educ. Psychol., 60, 10-22, 1990

INSTABILITY OF PERFORMANCE ON COGNITIVE TESTS IN INFANTS AND YOUNG CHILDREN WITH DOWN'S SYNDROME

BY JENNIFER G. WISHART AND LOUISE DUFFY
(Edinburgh Centre for Research in Child Development)

SUMMARY. Stability of cognitive performance in assessment situations was investigated in a group of 18 children with Down's Syndrome (DS) aged 6 months to 4 years. Two widely used tests of early cognitive development were presented: the Mental Scale of the Bayley Scales of Infant Development and a series of Piagetian object concept tasks. Both sets of tasks were presented in the same testing session, with testing repeated one to two weeks later. Both qualitative and quantitative aspects of performance were recorded. Even when overall scores were similar, patterns of pass/fails were very different in the two testing sessions. Successes showed poor stability over sessions with fails often occurring by default, the result of a refusal to engage fully in a particular task. This pattern of results suggests: (1) that single-session testing may not adequately assess cognitive status in DS children; (2) that the relationship between performance and competence may be unstable in such children, and (3) that "slow development" theories may not adequately describe cognitive development in young children with DS.

INTRODUCTION

WHILE considerable number of children with Down's Syndrome (DS) may also have motor or sensory handicaps, the major handicap is clearly cognitive. DS is the most common form of mental handicap in the UK and world-wide, accounting for approximately one-third of all children with severe mental handicap. DS provides researchers with a unique opportunity to investigate developmental processes in handicap: unlike most other forms of mental handicap, it occurs in large numbers, it is of known genetic origin and, most importantly, it can be confidently identified at birth. Careful, detailed study of early learning ability in children with DS may thus present us with unique insights into developmental processes in handicap.

One of the major debates in research on development in mentally handicapped children is whether developmental processes are broadly equivalent to those seen in normal development (only taking longer to unfold and progressing to a lower ceiling) or whether there are significant differences in both structure and organisation — the "difference versus delay" debate (Zigler and Balla, 1982). Recently, a growing body of evidence has suggested that learning in children with mental handicap may indeed differ in very fundamental ways from that seen in non-handicapped children, with important qualitative as well as quantitative differences existing between the two populations (Rogers, 1977; Morss, 1983, 1985; Macpherson, 1984; Rondal, 1984, 1988; Cherkes-Julkowski *et al.*, 1986; Wishart, 1986, 1987; Duffy and Wishart, 1987; Moss and Hogg, 1987; Dyer *et al.*, 1988). If this is the case, both developmental theory and educational practice will require substantial revision.

Despite the increasing evidence that a delay theory may not adequately describe developmental processes in handicap, a surprisingly large number of studies have prejudged the issue in their experimental design, either explicitly or implicitly. This is particularly obvious in studies in which tasks designed for and standardised on young, non-handicapped children are presented to older, handicapped children. Failure is typically interpreted as evidence that a skill has yet to be acquired, that there is still a basic lack of competence in the area under study. Without evidence of the suitability of such tasks for children of that age, whether handicapped or non-handicapped and in the absence of developmental data on young handicapped children on these same tasks, such conclusions cannot be justified. The tasks may simply have failed to engage the older child, a case of "won't do" rather than "can't do" (Koegel and Mentis, 1985; Wishart, 1987).

The assumption of slow development may be less obvious but is equally prevalent in other types of study. The widespread practice of matching handicapped and non-handicapped children on the basis of mental age (MA), for example, also assumes that developmental processes are identical in both sets of children and that both have reached the same stage in development by the same developmental pathways. This assumption seems unwise. It is well known that two children with identical MAs may nonetheless have very different developmental profiles. The validity of MA scores as a measure of cognitive ability, moreover, depends on competence being accurately reflected in performance. Given the substantial and inevitable differences in learning histories, both in terms of length and experience of success/failure, it seems unlikely that performance and competence will be similarly linked in handicapped and non-handicapped children. This suggestion is backed up by findings from a number of studies which indicate that differences in performance between MA-matched handicapped and non-handicapped children are often due to differences in motivation, the handicapped child frequently underperforming in many situations (Snyder, 1977; Siegel, 1979).

Trying to decide between the relative merits of the delay and difference arguments may in fact turn out to be a fruitless pursuit, as impossible a task as trying to portion out the exact and relative contributions of genes and environment to development. To some extent, much of the argument depends on the level of analysis employed: studies in which the evidence appears to support a slow development theory may, on more detailed analysis, reveal the existence of importance differences (Kiernan, 1984). What it does draw our attention to, however, is the need for more detailed information on development in children with handicap, and especially for a better description of the earliest stages in that development. Until recently, surprisingly few data were available in sufficient detail to allow insightful comparison of developmental processes in handicapped and non-handicapped children. The majority of studies simply told us what we already know — that children with mental handicap do less well on almost any psychological test than do their non-handicapped peers.

A growing number of studies are now directed at filling in these gaps in our understanding. The authors are presently conducting a series of interlinked cross-sectional and longitudinal studies of early learning and development in infants and young children with DS. These in-depth studies aim not only to assess the deficits in learning but also to identify the processes which impede that learning. The longer term aim is to identify the learning "style" of the child with DS and from that, to generate effective teaching procedures (for an overview, see Wishart, 1989). In the particular study to be described here, two tests of cognitive ability in infancy were presented on two occasions to a group of 18 infants and young children with DS. The two tests used were the Mental Scale of the Bayley Scales of Infant Development (BSID) and a set of Piagetian object concept tasks (covering Piaget's Stages III-VI). These tests were chosen firstly, because a substantial data-base, including detailed norms and reliability measures, was already available for non-handicapped children (Werner and Bayley, 1966; Bayley, 1969; Uzgiris and Hunt, 1975; Wishart and Bower, 1984, 1985) and secondly, because both tests have been widely used in studies of mental development in handicapped children (for reviews see Gibson, 1978; Carr, 1985; Morss, 1985).

The study had three main aims:

(1) to investigate the stability of performance of infants and young children with DS on standardised tests of cognitive development,

(2) to investigate the relationship between performance and competence in children with DS, and

(3) to investigate the adequacy of describing early cognitive development in DS in terms of normal developmental processes.

These aims seemed important for practical and theoretical reasons. The theoretical issues have already been discussed above. On the more practical level, assessment of cognitive ability plays a far greater role in the life of young children with mental handicap than in that of ordinary children (Morss, 1988). Many decisions, most obviously those concerning educational placement, are based on outcome measures from cognitive tests such as the ones used here. More often than not, the test is given only once and generally by an adult who is unfamiliar with the child; frequently the mother is not present. If such tests do not accurately measure mental ability in handicapped children or do not do so reliably, re-evaluation of the value of such testing would be indicated. If, moreover, the inadequacy of these tests results from basic differences in the ways that children with handicap approach learning, a re-assessment of teaching methods would also be required: traditional teaching methods which are highly effective with ordinary children may be ineffective — or even counter-productive — when applied to mentally-handicapped children.

METHOD

Sample

The Mental Scale of the BSID is designed to test for abilities which normally emerge between birth and 30 months; object concept tasks cover a similar age range, birth to 24 months. Cognitive development in children with DS is generally described as progressing at approximately half the normal rate (see, e.g., Berry *et al.*, 1984). In order therefore to examine the validity and reliability of the two chosen tests over their whole range, 18 DS children in the age range 6 months to 4 years were recruited to the study, three at each of the following ages: 6, 12, 18, 24, 36 and 48 months.

All children were confirmed by karyotyping as having Down's Syndrome (DS): all were standard Trisomy 21, with one also showing a small percentage of XXX cells. There were 13 females and 5 males. Nine children had congenital heart disease of varying degrees of severity (the majority minor), three had diagnosed hearing loss, four had minor visual impairments and one suffered from arthritis; none, however, had been hospitalised for any lengthy period during their development. All children had been volunteered by their parents and came from a wide range of socio-economic backgrounds. Seven were first borns. Number of siblings at time of testing ranged between none and four.

None of the children was naive to assessment situations. All were being regularly visited at home by the local authority's home teaching service and all but the three youngest and one 2-year-old had already participated in studies carried out by the present investigators. This is clearly a possible source of confounding variance but is an unavoidable feature of carrying out a research programme with any clinical population of restricted size and with this population in particular. As the average interval since last being seen by the authors was three months, it was considered unlikely that this factor would seriously affect the validity of this study.

Each child in the present study served as his or her own control. Since both tests had already been standardised on large groups of non-handicapped children and since the specific purpose of the study was to investigate the validity and reliability of the two chosen tests within this particular population, no direct control group of non-DS children was considered necessary. As indicated in the introduction, the present authors have in any case severe reservations, both practical and theoretical, about the value of using control groups matched on the basis of mental age. It was hoped that the present study would provide additional empirical support for this viewpoint. Controls matched for chronological age also seemed to have little to recommend them, being likely only to confirm what is already known — that a child with DS will perform less well on cognitive tests than a child of the same age who has no handicap (Duffy and Wishart, 1987; Wishart, 1987).

Procedure

All testing took place in the Department of Psychology in a small, quiet room with minimal distracting features. Testing was scheduled for a time of day when the mother thought her child was most likely to be alert and co-operative. The mother was present throughout and the child sat on her knee for testing; mothers were asked to encourage their children if necessary but not to direct their responses in any way.

Two sets of tasks, widely used to assess early cognitive ability, were presented: a series of Piagetian object concept tasks (presented by the first author), followed by items from the Mental Scale of the BSID. Testing was repeated one to two weeks later, with procedure identical in both sessions. All testing sessions were videorecorded for subsequent analysis.

Object concept tasks: Object concept development is indexed by performance on a series of search tasks in which the hiding sequence increases in complexity according to the stage being tested (Piaget, 1937). In all but the simplest task, the infant must choose between one of two identical occluders in order to retrieve a hidden toy; unless he or she fully comprehends the particular hiding sequence, performance can only be at chance level.

Four levels of task, covering Stages III-VI of object concept development, were presented in ascending order of difficulty: Task 1, a simple 1-occluder hiding task; Task 2, a task involving retrieving an object from inside one of two possible occluders, the object having been visibly displaced from one to the other (widely known as the "AAB" task); Task 3, an inference task, a task in which, on discovering the absence of the object from an expected location, the child must infer its true location; and Task 4, a 2-occluder "switching" task in which, following the hiding of the object, the positions of the two occluders are transposed before the child is allowed to search. White cardboard cups served as occluders and a variety of small toys as objects. Four trials of each task were given, with side of hiding randomised over trials. Success on all four trials was required for accreditation of a pass on any given level of task.

Exact details of task procedure and a full justification of scoring criteria adopted can be found in Wishart and Bower (1984). In that study, the ages by which 75 per cent of infants could pass Tasks 1 to 4 were 5, 10, 15 and 22+ months respectively.

Mental Scale of Bayley Scales of Infant Development (BSID): Items from the BSID Mental Scale were presented in accordance with the procedural instructions given in the test manual (Bayley, 1969). A basal level was established and higher level test items then presented until a minimum of six consecutive failures had been recorded. From the raw score thus obtained, children were given an age equivalent (AE) level and, where possible, an MDI (Mental Development Index) score from the BSID normative tables.

RESULTS

Both qualitative and quantitative analyses of performance on the two task sets were carried out.

Object concept tasks

Success rates: Scores from the two testing sessions for each of the four levels of object concept task are shown in Table 1.

Success rates were not impressive at any age in either session. If total scores are considered, there was some evidence of scores increasing with increasing age (Jonckheere $S = 59$; $P<0.05$); this increase was very modest, however, given the norms for these tasks and the large difference in age between youngest and oldest children ($3^1/_2$ years). With the exception of Task 1, success rates were generally poor on all levels of task. The results from Task 2 were surprising. The age by which most non-handicapped children can pass this task is 10 months. As might have been expected, the majority of 6-12 month-old handicapped children failed in both sessions; all but one of the 18-24 month-olds passed both times. This relatively easy task appeared to present difficulties to the 3-4 year-olds, however, with four failing in at least one session. Task 3 showed a more usual developmental pattern, with young children failing and only the very oldest children succeeding. Task 4 seemed difficult at all ages, although there was some improvement in older children by the second testing session.

TABLE 1

PASS RATES ON EACH OF THE FOUR LEVELS OF OBJECT CONCEPT TASKS IN THE TWO TESTING SESSIONS

	Task														
	1			2			3			4				
Testing Session	I	II		I	II		I	II		I	II			
Age (months)													Pass rate (all tasks combined)	N Child varying over sessions
6	1	2	(1/0)	0	0	(0/0)	0	0	(0/0)	0	0	(0/0)	3	1
12	3	3	(0/0)	2	1	(1/2)	0	0	(0/0)	0	1	(1/0)	10	4
18	3	3	(0/0)	3	3	(0/0)	0	0	(0/0)	0	1	(1/0)	13	1
24	3	3	(0/0)	3	2	(0/1)	0	0	(0/0)	0	0	(0/0)	11	1
36	3	3	(0/0)	1	2	(1/0)	3	3	(0/0)	0	1	(1/0)	16	2
48	3	3	(0/0)	1	1	(0/0)	3	2	(0/1)	0	2	(2/0)	15	3
Totals:	16	17	(1/0)	10	9	(2/3)	6	5	(0/1)	0	5	(5/0)	68 (8/4)	12

Note: Numbers in parentheses: N children whose performance improved over sessions/ N children whose performance deteriorated.

Stability of pass/fails: Although the grouped data indicated that success rates varied very little between sessions, this was not the case when individual records were examined. The performance of 12/18 children differed on at least one task in the two testing sessions (see Table 1). In eight cases, performance improved while four children failed to repeat an earlier success on re-testing (one 12-month-old improved performance on Task 4 but failed at a previously-passed level, on Task 2).

Order of difficulty of tasks: Normative studies have established that these tasks form a developmental hierarchy, ascending in order of difficulty from Task 1 through to Task 4: a child who can pass Task 3, for example, can reliably be expected to pass Tasks 1 and 2 but not Task 4. Group scores for each of the four levels of task appeared to confirm this as the order of difficulty pertaining in the DS children tested here (see Table 1). Closer examination of individual patterns of pass/fails did not support this conclusion, however.

Differences in difficulty level between any two tasks on either testing session were computed for each child and the derived difficulty order compared with the normatively-derived sequence. As in any task order analysis of cross-sectional data, each child provided data on the relative difficulty of only a subset of the tasks presented, those tasks falling to either side of his or her current developmental stage; any child who either failed or passed all tasks (two in this study) provided no relevant data. (For example: a child who passed Tasks 1, 2 and 4 but failed Task 3, provided no data on the relative difficulty of Tasks 1, 2 and 4 to each other but did provide data on the difficulty of Task 3 in relation to each of these tasks; since, moreover, Task 3 was found to be more difficult than Task 4, the derived difficulty order in this case could be classed as deviating from the normatively-derived hierarchy for these tasks.)

In six out of the 16 cases where comparative information on task difficulty existed, order of difficulty did not match the norm for at least one task pairing ($\chi^2 = 6.58$ (df1): $P<0.01$). In two younger children, this was due to a reversal in difficulty level between Tasks 3 and 4; in four older children, it was due to Task 2 proving more difficult than Task 3.

Failures to engage: Over the two testing sessions, there were 76 instances where a child's performance did not meet the stringent statistical criteria necessary to score a pass. Inspection of performance protocols, however, suggested that many of these failures could not confidently be attributed to a straightforward absence of the required level of cognitive ability. In 55 cases, there was evidence that the child had not fully engaged in all of the four trials presented (Table 2), a frequency which would suggest extreme caution in using such data as a basis for assessing cognitive status in these children.

TABLE 2

FREQUENCIES OF FAILURES TO ENGAGE ON EACH OF THE FOUR OBJECT CONCEPT TASKS IN EACH AGE GROUP

Age (months)	Session No.	1	2	3	4	Totals
		t/p	t/p	t/p	t/p	t/p
6	I	1/0	2/0	2/0	2/0	7/0
	II	0/0	2/0	1/1	2/0	5/1
12	I	0/0	1/0	1/2	2/0	4/2
	II	0/0	0/1	0/3	0/1	0/5
18	I	0/0	0/1	1/1	0/0	1/2
	II	0/0	0/0	1/2	0/1	1/3
24	I	0/0	0/0	1/1	1/0	2/1
	II	0/0	1/0	1/1	1/0	3/1
36	I	0/0	1/2	0/0	3/0	4/2
	II	0/0	0/2	0/0	1/2	1/4
48	I	0/0	1/1	0/0	0/1	1/2
	II	0/0	1/1	1/0	1/0	3/1
Totals:		1/0	9/8	9/11	13/5	32/24

Note: t = total failure to engage; p = partial failure to engage
(see text for definitions)

Failure to engage occurred with near equal frequency in both sessions. It was defined as any instance where a child refused to attend to the hiding sequence or failed to search quickly and clearly for the hidden object (the success or failure of any search being immaterial). Failure to engage was classified as being either total or partial. It was classified as total if on any trial, the subject either (a) refused to attend to the full hiding sequence or (b) made no attempt at search. Since the criterion for a pass on all tasks was 4/4, this level of "switching out" meant that some children failed by default, rather than by making errors. Interestingly, many such children would happily re-engage, albeit it often only temporarily, in subsequent trials of another task, even although that task involved exactly the same test materials and was of a higher level of difficulty than that just refused. Also noteworthy was that all but the very youngest children were as likely to opt out after a trial in which they had successfully retrieved the object as after one in which initial search had been unsuccessful.

Failures to engage were often only partial. Although prepared to watch the hiding sequence and to search, it was clear that that search did not represent the child's best attempt at the problem being presented. As trials progressed, children would delay search, often attempting to divert the tester into some sort of social game by producing some sort of party-trick (such as clapping hands) or response which was quite unrelated to the task in hand. This sort of behaviour can only make a difficult task more difficult, adding a memory component where none previously existed. Some children simply opted out of a particular task by refusing to commit themselves to either cup, sweeping both from the table, an inefficient, low-level strategy which led to recovery of the object but did not meet the criteria for a pass.

Sided strategies: Less readily classifiable as refusals to engage (and not included in Table 2) were instances where children adopted a sided strategy, automatically choosing the cup on one side on every trial of a given task, irrespective of whether that cup had actually been involved in the relevant hiding sequence or not. There were 14 instances of a sided strategy being used, all but one occurring in relation to the two higher level tasks, Tasks 3 and 4.

These rates do not differ significantly from the frequencies of usage that would be expected on the basis of normative studies of these tasks. However, the quality of the behaviour and the response to its outcome were noticeably different to the first author, a highly experienced tester both of these tasks and of this population. A sided strategy is "correct" as often as it is "wrong" in the design adopted here.

Typically, however, on trials where the strategy in fact failed, children showed no surprise or concern, simply going on quickly to search at the remaining cup. The side repeatedly chosen, moreover, corresponded to the child's demonstrated hand preference in all but two instances. This is not unusual for Task 4 but is highly unusual for Task 3, where a sided strategy is generally the result of a tendency to repeat search to the side which first produced success. This suggests that while on the surface these children appeared to be co-operating, there was little true task engagement, the strategy adopted being one of convenience rather than one evolved from cognitive considerations.

A number of other indications of falling levels of task engagement were repeatedly observed within trial sets of a given task. Attention became increasingly difficult to capture and often drifted during hiding sequences, and search was clearly less motivated. In Task 3, for example, exhaustive search of all possible locations in early trials fell to 1-stage search in later trials: both qualify as "fails" on statistical criteria but the number of locations searched in this task has been shown in normative research to be a good indicator of developmental status. Similar evidence of reduced effort was found for Task 4; errors — quickly corrected in initial trials — were simply accepted in later trials, the child sitting, as if helpless, waiting to be shown where the toy was. Since introduction of a new level of task led to a revival of effort and attention, this lower-level responding could not be attributed to simple fatigue effects, to basic attentional deficits or to an overall low level of motivation.

Mental Scale of Bayley Scales of Infant Development (BSID)

Raw scores: In over half the testing sessions, raw scores on the BSID Mental Scale were too low to be converted into Mental Development Indices (MDIs). BSID data were therefore treated in terms of raw scores and their age-equivalent (AE) levels.

Table 3 shows the mean raw scores in the two testing sessions for the six age groups tested. Raw scores increased significantly with increasing age (Jonckheere $S = 127$; $P<0.01$ for both sessions). Even the six-month-olds, however, scored 1 to 2 months below the norm for their age, a deficit in performance which widened with increasing age, the oldest children scoring at levels between 18 and 28 months behind expected levels for non-handicapped children of similar age.

TABLE 3
RAW SCORES AND VARIABILITY ON THE BSID MENTAL SCALE IN THE SIX AGE GROUPS

Age (months)	Testing Session I	Testing Session II	Total N items varying over sessions	
	Mean score (AE)		Fail-to-pass	Pass-to-fail
6	55.3 (4.5)	53.3 (4.5)	12	19
12	77.3 (7.0)	75.6 (7.0)	10	15
18	100.0 (11.0)	99.7 (11.0)	11	13
24	109.0 (13.0)	108.7 (13.0)	16	16
36	131.3 (19.0)	130.3 (19.0)	12	15
48	142.3 (22.0)	142.3 (22.0)	13	13
Totals:			74	91

Note: AE = age equivalent level (in months)

Stability of pass/fails: In order to examine the stability of pass/fails, two measures of reliability were taken: a comparison of the raw scores achieved by each child in the two testing sessions (the most direct measure of the reliability of an "intelligence" test) and a more detailed analysis of the item-item consistency for each child over the two testing sessions. Quite different pictures emerged from these two analyses. Comparison of raw scores indicated that there was no significant difference in performance over the two sessions (t = 0.76, (df 17); NS). Examination of individual performance profiles, however, showed that even where raw scores were very similar (or even identical), performance had in fact changed on a substantial number of items (see Table 3).

Item-item agreement over sessions in individual children would be the most obvious and sensitive measure of a test's reliability. No such normative data are provided by Bayley, however. Re-test data from a very small sample of children (28) at only one age level (8 months) are presented and no indication of the reliability of test items for individual children is given, only the relative reliability of individual test items. Procedure, moreover, varied from the standard testing procedure in that all re-test children were given the same number of items, irrespective of their individual ability levels. In order, therefore, to investigate whether variability in performance was greater in the sample of DS children tested here than in the normative sample, the variability in each individual child's scores over the two testing sessions was compared to the variability which would be predicted on the basis of the standard errors of measurement (SEm) provided in the BSID manual for the range of mental ages covered by the test. As a measure of reliability, this is likely to be an over- rather than under-estimate of the variability in scores to be expected in individual children.

In most cases, scores were too low to be converted into MDIs. Comparisons were therefore carried out against the variability predicted in raw scores on the basis of the SEm and AE level for each child (for a child of 24 months with an age equivalent level of 12 months, for example, raw scores would be expected to vary by a maximum of 6 points on re-testing). A t-test indicated that there was a significant difference between observed and expected variability over the two sessions ($t = 2.78$, df (17), $P<0.01$). Overall, there were 165 cases of item instability: in 74 cases, items which had been failed in the first session were passed in the second session; in 91 cases, the change was in the opposite direction, a previously passed item being failed on re-testing ($\chi^2 = 0.96$ (df 1): NS). Degree of variability and ratio of pass-to-fail to fail-to-pass changes were similar at all ages ($t = -0.82$, (df17); NS).

Patterns of fails-to-passes and of passes-to-fails are equally interesting but for different reasons. A fail-to-pass could result from a genuine developmental advance having been made between sessions. However, given the short interval between testing sessions, this was unlikely to be the case when performance improved on a substantial number of items in the second session. It was also unlikely that any such improvement could have resulted merely from increased familiarity with the tester on the second session, since she was already familiar to most of the children on the first testing session. More likely is that fluctuating levels of task engagement were responsible for the instability shown. In this respect, a pass-to-fail pattern is far less ambiguous. Clearly the required behaviour *was* in the child's repertoire but was not for some reason being reproduced on the second testing. This instability must inevitably lead to underestimation of the child's true competence. Unlike fails, any pass, whether produced in the first or the second session, is likely to be genuine; few of the BSID items permit chance success. When the children tested here were credited with all items on which they had scored a pass, regardless of its stability, this "optimal" score proved — not surprisingly — to be consistently higher than the score actually achieved in either of the two testing sessions (1st/2nd session: $t = 2.60/3.24$ (df17); $P<0.01/0.005$).

A further analysis of the difficulty level of those items which proved to be unstable was carried out in which each item which varied was related to the developmental level of each child. The midpoint between lowest basal and highest ceiling levels attained over the two sessions was determined for each child and the numbers of pass-to-fail and fail-to-pass items which fell below and above this midpoint were calculated. Numbers of fail-to-pass items fell equally on either side of this midpoint. There was, however, a trend for more pass-to-fail items to occur above than below this midline and for these to exceed the number of fail-to-pass items above this point ($t = 1.40/1.24$ (df17); $P<0.10$/NS respectively).

Failures to engage: Given how unstable performance was, it seemed worth investigating the possible source of this instability. Failures to engage were more difficult to define in relation to the BSID items. Unlike the object concept tasks, BSID items vary greatly, both in content and in terms of task demands. It is particularly difficult to identify with any confidence any failure to engage in relation to items at the bottom end of the scale, where to be credited with a pass the child has simply to respond to a stimulus (e.g., regards cube, turns head to rattle). In these cases, it is impossible to distinguish between a criterion fail, indicating absence of the required level of cognitive ability, and a fail by default, where failure is due to the child having refused to engage. At all levels, the BSID also includes a number of "incidental" items, items in which there is no particular task to be engaged in (e.g., vocalises once or twice, repeats performance laughed at); by definition, these too cannot be incorporated into any failure-to-engage analysis. Forty-seven out of the 165 cases of item instability were omitted from the analysis below on these grounds, although it is perhaps worth noting that as performance did differ on these items (and not in all cases in the fail-to-pass direction), it seems highly probable that something other than lack of the required cognitive ability was determining poor performance in one or other session.

In the 118 remaining eligible cases, subject records were re-examined for any evidence of failure to engage on the failed item. Failure to engage was defined as any instance where the child's response fell into one of the following three categories: (1) refusal to attend to the task, (2) rejection of the task object by casting, swiping or dropping it on the floor, or (3) repeated production of an inappropriate, off-task behaviour in response to the task materials. In over half of the cases where performance on items differed over sessions, failure was clearly due to the child refusing to engage in the task: in 40/64 cases in the pass-to-fail direction, and in 26/54 in the fail-to-pass direction.

DISCUSSION

The low levels of ability demonstrated were somewhat discouraging, given the effort and funds presently being diverted into early intervention programmes for DS children. Unlike DS children in the 1960s, all of the children in this study had home teachers, all were being brought up in their own homes and all had highly motivated parents, all of whom believed in actively encouraging their child's development in all areas. Ability levels differed very little, however, from those seen in earlier studies of this population (for overview, see Gibson, 1978). On average, performance was at approximately half the level that would have been expected given their chronological ages, a finding consistent with recent results from studies in the USA, Australia, Canada and Germany (Pueschel, 1984; Rauh *et al.*, 1987).

Although at one level the results lend support to a "slow development" theory, there did seem to be important differences in the way in which the children with DS responded to both sets of tasks. These particular tasks were selected for study because although neither test was specifically designed for or standardised on handicapped children, both have been used widely to assess early cognitive development in handicapped children. Object concept development was considered by Piaget to be prototypical of all cognitive development, and fifty years on object concept studies still dominate much of the literature on learning in handicapped and non-handicapped children (for reviews, see Schuberth, 1982; Harris, 1988). In a review of psychological tests of early cognitive ability, Francis *et al.* (1987) described the Bayley Scales as "unrivalled at determining a child's developmental status in relation to its age-mates". Only a few studies suggest that either test has any great predictive validity (e.g., Wachs, 1975; Siegel, 1981), but it is generally accepted that each has both concurrent and construct validity — that is, that they reliably measure in a single testing session whether a child has or has not reached important stages in early cognitive development.

The present results would suggest that neither test does this in the case of mentally handicapped children. On both sets of tasks, the variability in performance was far greater than would have been expected on the basis of large-scale normative studies (Bayley, 1969; Horner, 1980; Morss, 1985; Wishart and Bower, 1985). In the object concept tasks, there was no simple pattern to the variability demonstrated, either in terms of age or task level. The performance of 12 children varied over sessions and all levels of tasks showed some instability (although it was most frequently seen in Tasks 2 and 4). "Switching out", however, was consistently associated with *all* of these apparent reversals in ability, children failing to complete all four trials or responding in an inappropriate or stereotyped fashion in either the first or second testing session. Failures to engage were equally prevalent in response to the Bayley items and it was clear from the analysis of "optimal" versus typical performance that in these tasks too, virtually all the children were underperforming to some degree in both sessions.

The many failures to engage suggest that motivational factors play an important role in the expression of competence in performance in young children with DS. In the object concept tasks, this "switching out" may have been responsible for the disruption found to the order of difficulty these tasks normally hold to each other. Longitudinal research with 3-5 year-olds and with infants with DS has indicated that failure on object concept tasks at later ages may not, as is commonly concluded, represent a failure to have acquired the requisite cognitive skills but may result, rather, from a failure to consolidate these skills when they were first acquired, in early infancy (Wishart, 1987; 1988a). In these studies, failure to engage appeared to result from the operation of two distinct processes, both of them linked to the relationship between the difficulty level of the task being presented and the child's stage in development at time of testing. Infants in these longitudinal studies went to considerable lengths to avoid tasks of a level of difficulty one step or more above their current developmental status. This was perhaps not surprising but unfortunately tasks previously mastered also fell outwith the range of engagement, performance typically falling away in the months following initial acquisition. Only tasks falling within a narrowly defined range of difficulty were fully engaged in and reliably passed, suggesting that this form of cognitive avoidance adversely affects both the acquisition and consolidation stages of learning. The analysis of the Bayley data lends some support to this interpretation, with "switching out" and instability occurring in relation to items at both the easy and difficult ends of the child's developmental range.

In both these previous Wishart studies and in the study reported here many task failures were by default, the result of the child refusing to participate in all trials of the task being presented. Often the child used — or more accurately, *mis*-used — social skills to get out of the task in hand (Wishart, 1988b). To classify such responses as "fails" or to describe the child as "untestable" is not particularly helpful, either to the child or to our understanding of that child. It is clear, however, that any attempt to modify performance in such situations will require a better understanding of developmental processes in children with handicap than that presently held. We shall also need deeper insight into the interrelationship of motivational factors, competence and performance in such children. In the case of some failures, it may be possible to modify testing procedures to suit the child's cognitive "style" better; it may be that the number of trials dictated by statistical considerations exceeds that considered to be cognitively or motivationally relevant to the post-acquisition child with DS. Overcoming poor engagement in the acquisition stage may be more difficult but it is perhaps possible that artificial enhancement of success/failure ratios could prevent avoidance behaviour becoming the routine response to difficult learning situations (Seligman, 1975).

Obviously, a cross-sectional study such as this cannot provide any direct evidence on development or developmental processes and cannot therefore explain the mismatch between performance and competence found in so many of the children here. The instability shown could result from poor motivation, could reflect a basic instability in the learning process itself or could be due to a complex developmental interaction between inadequate motivation and inefficient learning processes, with poor motivation contributing directly to the growth of deficits in cognitive functioning. What can safely be concluded is that, if the instability demonstrated in assessment situations is characteristic of the DS child's response to his or her everyday learning environment, quite clearly inefficient use is being made of whatever level of ability is present, a factor which can only add to the child's already existing handicap.

Unstable performance may underlie a feature frequently commented on in relation to development in DS. Parents of DS children often comment that their child's development seems to be very uneven, several advances being made in a comparatively short period of time, only to be followed by a lengthy period in which development seems to have come to a virtual standstill. A number of researchers have found evidence supporting the suggestion that development in DS shows such "plateaux", periods which they have interpreted as periods of consolidation of recently acquired skills (e.g., Carr, 1975; Cunningham, 1988). Others are less optimistic, seeing these plateaux as evidence of the child meeting a "developmental wall", a period during which no development takes place at all (Gibson, 1978). Neither may in fact be an accurate description of what is going on. If individual profiles are monitored, we may find that these apparent lulls in development are as much due to the dropping out of previously passed items as to any failure to make progress on new, higher level tasks.

A longitudinal study presently in progress aims to test out this alternate hypothesis. If this does show poor stability of new acquisitions, clearly teaching effort will have to be directed not only at encouraging development of new skills but also at ensuring that recently acquired skills are adequately consolidated. How easy it will be to bypass the DS child's apparent inclination to avoid this remains to be seen. Some evidence does exist that tailoring of teaching methods to the particular needs and skills of the handicapped child can improve performance and can lead to positive carry over effects in less structured learning situations (see, e.g., Duffy and Wishart, 1987); it is not yet clear, however, whether any long-term benefits accrue from this type of approach (Wishart, 1986). Amplified developmental instability may, as Shapiro (1983) suggests, be built into the learning system, an inevitable effect of the trisomy on development, affecting all but the most buffered areas of development; the degree to which this instability may be compensated for through environmentally-based intervention remains an open question.

ACKNOWLEDGMENTS.—This research was supported by MRC Project No. G8703875N and by ESRC Linked Studentship Award No. C00428725008. We would like to acknowledge, with very many thanks, the help of all of the mothers and infants who took part in this study. We are also grateful to the Scottish Down's Syndrome Association for the constant encouragement and support which they have given to us throughout our research programme. Thanks are also extended to the Department of Child Life and Health, University of Edinburgh, to the Cytogenetics Lab, Royal Infirmary for Sick Children (Edinburgh) and to the Department of Human Genetics, Western General Hospital (Edinburgh) for providing karyotype information.

Correspondence and requests for reprints should be addressed to Dr Jennifer Wishart, Edinburgh Centre for Research in Child Development, Department of Psychology, 7 George Square, Edinburgh EH8 9JZ, Scotland.

REFERENCES

Bayley, N. (1969). *Bayley Scales of Infant Development*. New York: Psychological Corporation.

Berry, P., Gunn, V. P., and Andrews, R. J. (1984). Development of Down's Syndrome children from birth to five years. In Berg, J. M. (Ed.), *Perspectives and Progress in Mental Retardation*. Baltimore: University Park Press.

Carr, J. (1975). *Young Children with Down's Syndrome*. London: Butterworth.

Carr, J. (1985). The development of intelligence. In Lane, D., and Stratford, B. (Eds.), *Current Approaches to Down's Syndrome*. London: Holt, Rinehart and Winston.

Cherkes-Julkowski, M., Gertner, R., and Norlander, K. (1986). Differences in cognitive processes among handicapped and average children: a group learning approach. *J. Learn. Disabil.*, 19, 438-455.

Cunningham, C. (1988). *Down's Syndrome: An Introduction for Parents. (2nd ed.)*. London: Souvenir Press.

Duffy, L., and Wishart, J. G. (1987). A comparison of two procedures for teaching discrimination to Down's Syndrome and normal children. *Br. J. educ. Psychol* , 57, 265-278.

Dyer, S., Gunn, P., Rauh, H., and Berry, P. (1988). Motor development in Down Syndrome children: an analysis of the Motor Scale of the Bayley Scales of Infant Development. Paper presented at the 8th IASSMD Congress, Dublin, Ireland, August.

Francis, P. L., Self, P. A., and Horowitz, F. D. (1987). The behavioural assessment of the neonate: an overview. In Osofsky, J. D. (Ed.), *Handbook of Infant Development (2nd ed.)*. New York: Wiley.

Gibson, D. (1978). *Down's Syndrome: The Psychology of Mongolism*. Cambridge: Cambridge University Press.

Harris, P. L. (1988). The development of search. In Salapatek, P., and Cohen, L. B. (Eds.), *Handbook of Infant Perception*. New York: Academic Press.

Horner, T. M. (1980). Test-retest and home-clinic characteristics of the Bayley Scales of Infant Development in nine- and fifteen-month-old infants. *Child Dev.*, 52, 751-758.

Kiernan, C. C. (1984). The development of communication and cognition. In Dobbing, J. *et al*. (Eds.), *Scientific Studies in Mental Retardation*, London: Macmillan Press.

Koegel, R. L., and Mentis, M. (1985). Motivation in childhood autism: can they or won't they? *J. Child Psychol. Psychiat.*, 26, 185-191.

Macpherson, F. D. (1984). The structure of sensori-motor intelligence in severely and profoundly mentally handicapped children. Unpublished PhD thesis, University of Southampton, England.

Morss, J. R. (1983). Cognitive development in the Down's Syndrome infant: slow or different? *Br. J. educ. Psychol.*, 53, 40-47.

Morss, J. R. (1985). Early cognitive development: difference or delay? In Lane, D., and Stratford, B. (Eds.) *Current Approaches to Down's Syndrome*. London: Holt, Rinehart and Winston.

Morss, J. R. (1988). Cognitive motivation in the young child with Down's Syndrome. Paper presented at the 8th IASSMD Congress, Dublin, Ireland, August.

Moss, S., and Hogg, J. (1987). The integration of manipulative movements in children with Down's Syndrome and their non-handicapped peers. *Human Movement Science*, 6, 67-99.

Piaget, J. (1955). *The Construction of Reality*. London: Routledge and Kegan Paul. (Original French edition, 1937).

Pueschel, S. M. (1984). *The Young Child with Down Syndrome*. New York: Human Sciences Press.

P, 7207

RAUH, H., BERRY, P., AND RUDINGER, G. (1987). Patterns of development in Down Syndrome children using the Bayley Scales. Paper presented at the Ninth Biennial Meeting of ISSBD, Tokyo, Japan, July.

ROGERS, S. J. (1977). Characteristics of the cognitive development of profoundly retarded children. *Child Dev.*, 48, 837-843.

RONDAL, J. A. (1984). Linguistic and pre-linguistic development in moderate and severe mental retardation. In DOBBING, J. *et al.* (Eds.) *Scientific Studies in Mental Retardation.* London: Macmillan Press.

RONDAL, J. A. (1988). Language development in Down's Syndrome: a life-span perspective. *Int. J. behav. Dev.*, 11, 21-36.

SCHUBERTH, R. E. (1982). The infant's search for objects: alternatives to Piaget's theory of object concept development. In LIPSITT, L. P., and ROVEE-COLLIER, C. K. (Eds.), *Advances in Infancy Research: Vol. 2.* Norwood, N. J.: Ablex.

SELIGMAN, M. E. P. (1975). *Helplessness.* San Fransisco: W. H. Freeman.

SIEGEL, P. S. (1979). Incentive motivation and the mentally retarded person. In ELLIS, N. P. (Ed.), *Handbook of Mental Deficiency: Psychological Research and Theory.* Hillsdale, N.J.: Erlbaum.

SIEGEL, L. (1981). Infant tests as predictors of cognitive and language development. *Child Dev.*, 52, 545-557.

SHAPIRO, B. L. (1983). Down Syndrome — a disruption of homeostasis. *Am. J. med Genet.*, 14, 241-269.

SNYDER, L. S. (1977). Communicative and cognitive abilities and disabilities in the sensori-motor period. *Merrill Palmer Q.*, 24, 161-180.

UZGIRIS, I., AND HUNT, J. McV. (1975). *Assessment in Infancy: Ordinal Scales of Psychological Development.* Urbana, Ill.: University Press.

WACHS, T. D. (1975). Relation of infants' performance on Piaget scales between 12 and 24 months and their Stanford-Binet performance at 31 months. *Child Dev.*, 46, 929-935.

WERNER, E. E., AND BAYLEY, N. (1966). The reliability of Bayley's Revised Scale of Mental and Motor Development during the first year of life. *Child Dev.*, 37, 39-50.

WISHART, J. G. (1986). The effects of step-by-step training on cognitive performance in infants with Down's Syndrome. *J. ment Defic. Res.*, 30, 233-250.

WISHART, J. G. (1987). Performance of young non-retarded children and children with Down Syndrome on Piagetian infant search tasks. *Am. J. ment. Defic*, 92, 169-177.

WISHART, J. G. (1988a). Early learning in infants and young children with Down's Syndrome. In NADEL, L. (Ed.), *The Psychobiology of Down Syndrome.* Boston: M.I.T. Press.

WISHART, J. G. (1988b). Misuse of social skills in young children with Down's Syndrome. Paper presented at the summer meeting of the Experimental Psychology Society, Edinburgh, Scotland, July.

WISHART, J. G. (1989). Learning to learn: the difficulties faced by infants and young children with Down's Syndrome. In FRASER, W. I. (Ed.), *Key Issues in Research in Mental Retardation.* London: Routledge (in press).

WISHART, J. G., AND BOWER, T. G. R. (1984). Spatial relations and the object concept: a normative study. In LIPSITT, L. P., and ROVEE-COLLIER, C. K. (Eds.), *Advances in Infancy Research: Vol. III.* Norwood, N.J.: Ablex.

WISHART, J. G., AND BOWER, T. G. R. (1985). A longitudinal study of the development of the object concept. *Br J. dev. Psychol.*, 3, 243-258.

ZIGLER, E., AND BALLA, D. (Eds.) (1982). *Mental Retardation: The Developmental Difference Controversy.* Hillsdale, N.J.: Erlbaum.

(*Manuscript received* 3rd *March*, 1989)

Br. J. educ. Psychol., 60, 23-36, 1990

OCCUPATIONAL OUTCOMES OF TRUANCY

BY ANGELIKA HIBBETT*
(Employment Department, London)
KEN FOGELMAN
(School of Education, University of Leicester)
AND ORLY MANOR
(Department of Statistics, University of Jerusalem)

SUMMARY. This paper presents some findings of a study into early adult outcomes of truancy, which was based on the National Child Development Study, a longitudinal survey of all people in Great Britain born in the week 3rd-9th March 1958. It was found that truancy is associated with lower status occupations, less stable career patterns and more unemployment. Among those who were working, former truants' incomes were not lower, but they were considerably less well off once their family situation was taken into account. Differences remained after controlling for the effects of social background, educational ability, poor attendance due to other reasons, and end-of-school qualifications. The authors conclude that truancy is a predictor of employment problems, and of a more severe kind than will be experienced by others who share the disadvantaged background and low attainment which typify the truant.

INTRODUCTION

TRUANCY, particularly in the final years of compulsory secondary schooling has been identified as a substantial phenomenon. However, both the underlying causes and the implications for the young people concerned are less well understood. Thus, the purpose of the study described here has been not only to provide descriptive information on how former truants were faring later in their lives, but thereby to shed some light on possible explanations. For example, if truancy is part of a more comprehensive maladjustment (Tyerman, 1968), this would be consistent with subsequent problems in early adulthood, in areas such as family relationships and mental health. If, as others have suggested (Haigh, 1976), truancy represents a move away from the values of the school and towards those of adulthood, then we should not expect former truants to encounter particular difficulties in their labour market experiences.

Outcomes of truancy appears to be a relatively under-researched topic, and such surveys as have been carried out have yielded some conflicting findings. In a study of causes and effects of persistent job changing, Cherry established low attendance as a good predictor of a subsequent unstable job record, but apart from less job satisfaction no occupational disadvantages were associated with persistent job changing and, in particular, men who frequently changed jobs earned as much as stable workers (Cherry, 1976).

Subsequently, in a longitudinal survey from age 8 to age 21 of 411 males in a working class area in London, Farrington (1980) found that truants were more likely to have lower status jobs and unstable job histories after leaving school.

Similar findings were reported by Robins and Ratcliffe (1980), who investigated 235 black schoolboys from St Louis, born between 1930 and 1934. In this study, truancy was inferred from poor attendance as revealed in school registers, and it is thus probably more appropriate to refer to persistent absenteeism rather than truancy in this context. Persistently poor attendance was associated with dropping out of secondary education, and low earnings as an adult. Earnings were predicted by absenteeism even when controlling for the effects of non-graduation: absentees who did graduate from high school had lower earnings than good attenders who graduated, although their earnings were higher than those of absentees who dropped out.

---

* This research was carried out when Dr Hibbett was employed at the Social Statistics Research Unit, City University, from 1986 to 1988.

Gray *et al.* (1980) looked at school attendance and the first year of employment, and found that absentees tended to leave school without formal qualifications. They were subsequently likely to be in less skilled work, had an increased risk of unemployment, and a reduced chance of being in a job involving further training. This adverse effect of absenteeism on the first year of employment operated through low scholastic achievements rather than other, for example psychological, mechanisms. Poor attenders had lower scholastic achievements, even when IQ, social class and other background variables were controlled for: for any given level of intelligence, possession of educational qualifications increased the likelihood of obtaining skilled work involving further training.

While there is thus some evidence to suggest that truancy and absenteeism are associated with certain subsequent occupational disadvantages, this is by no means a generally observed trend; more particularly, the findings of Cherry contradict those of Robins and Ratcliffe regarding lower earnings as an adult. A tendency to have less skilled jobs and an unstable job history, as results of low scholastic achievements and lack of qualifications, appears to be generally established.

Three main issues underlie the analyses reported in this paper. Firstly, as reported below, preliminary analyses demonstrated a consistent relationship between truancy and virtually all the outcomes examined. This, however, is the result of comparing the truants with all other members of the 1958 Cohort, and earlier work (Fogelman and Richardson, 1974) had established that truants differ from other young people in a number of respects which might be expected to explain such relationships. For example, they came predominantly from families where the fathers were in manual occupations, they were likely to be in larger families, they were not equally distributed in different parts of the country, and they showed poorer educational ability and attainment even before they were identified as truanting. Thus the first stages of the analyses reported here are designed to assess whether taking such differences into account eliminates or reduces the relationship between truancy and outcomes. In other words, they investigate whether the outcomes for truants differ once they are compared with people who are similar in these respects.

Secondly, although there are other reasons for being absent from school, it seems likely — and the preliminary analyses confirmed — that truants have lower than average school attendance. It is known that poor school attendance, for whatever reason, is associated with lower attainment (Fogelman, 1978). Thus, one stage of the analysis is designed to establish whether any differences in outcomes for the truants (after allowing for the background factors as described above) are explained by their depressed attendance levels.

Thirdly, as noted above, one of the previous studies in this area (Gray *et al.*, 1980) suggested that difficulties encountered by truants, particularly in the labour market, were the result of their depressed examination performance and lack of qualifications at the end of their schooling; and that thereafter their problems were no greater than others with similarly poor qualifications. The final stages of our analyses address this issue.

These analyses are based on data from the National Child Development Study. The breadth of data available in this study and its longitudinal structure make it possible to examine the relationship between reported truancy from school, based on information obtained during the school years, and a number of areas explored during interview at the age of 23, including occupational, financial and educational progress and aspects of family formation and health-related behaviour. The current paper concentrates on occupational outcomes. A further paper is in preparation describing the other areas. A full account of the analyses and findings of our study is available in Hibbett (1987) and Hibbett and Fogelman (1988).

METHOD

The National Child Development Study (NCDS)

NCDS is a longitudinal, multi-disciplinary study of all the people in Great Britain who were born in the week 3rd-9th March, 1958. It has its origins in the 1958 Perinatal Mortality Survey, carried out by the National Birthday Trust Fund (Butler and Bonham, 1963). Subsequent follow-ups of the 15,000 or so survivors were carried out by the National Children's Bureau at the ages 7, 11, 16 and 23 (Davie *et al.*, 1972; Fogelman, 1983). In addition schools were approached in 1979 for the results of public examinations taken up to that time and there have been a number of special exercises to collect data on specific subgroups. Response has generally been high, at about 90 per cent during the school years and 76 per cent at age 23.

The information obtained during the school years included indicators of truancy, usually individual items from standard behaviour scales (Stott, 1969; Rutter *et al.*, 1970), completed by teachers, parents and the young people themselves.

A composite measure of truancy was created based on the teachers' ratings, because of their more consistent relationship with attendance rates (Hibbett, 1987). The truancy measure identified those said to be truanting at 16, and at both 11 and 16, and the comparison group who were not said to be truanting at either age. It was apparent that the 'Truanting at 11 and 16' group was too small to be included in the eventual multivariate analysis, but it was felt useful to continue to explore their simple associations. This was not the case for the very small group identified as truanting at age 11 but not at 16 (Table 1).

TABLE 1

NUMBER OF RESPONDENTS

Number of Cases Column %	Boys	Girls	Total
Truanting at 11 only	26	15	41
	0.5	0.3	0.4
Truanting at 11 and 16	51	11	62
	0.9	0.2	0.6
Truanting at 16 only	1179	888	2067
	21.7	17.1	19.4
Not truanting at 11 and 16	4178	4292	8470
	76.9	82.4	79.6
Column Total	5434	5206	10640

In this and the following tables, there is a sizeable proportion of missing values, reducing the number of cases included in the analyses. Information on truancy is available for 10,640 respondents, which constitutes 85 per cent of the total sample of 12,537 respondents at age 23. Analysis of various characteristics leads to further attrition in the sample, resulting from cumulative non-response, which varied according to the outcomes analysed. Most of this non-response results from failure to trace, refusal, or failure for administrative, resource or other reasons to get one or more survey instruments completed.

Where analyses, such as these, draw on data collected at several ages and by more than one instrument within any one follow-up, the effect is cumulative and can be substantial.

It is possible to make some assessment of the effects of non-response in a longitudinal study, as the earlier characteristics of subsequent non-respondents are known. General analyses of NCDS are reasonably reassuring. Respondents are slightly more middle class, and certain disadvantaged groups are becoming under-represented, but in general such biases are small. The more troubling bias is that ethnic minority groups are substantially under-represented (Iyer, 1984). Perhaps most relevantly, it has been demonstrated that repeating earlier analyses omitting subsequent non-respondents has no more than trivial effects on the results (Goldstein, 1983).

More specifically for the present analyses it is important to examine the characteristics in the reduced sample with data on all the variables included in the multivariate analyses. This is done in two ways. Firstly, it is possible to compare the unadjusted results in the tables presented in this paper with the simple two-way tables which require data on only the truancy variable and the outcome. Secondly, for a selection of outcomes, we have recalculated the earlier stages of the multivariate analyses without omitting the cases which were excluded because of missing data on variables included in later stages of the analysis. In neither case did this indicate major differences — certainly not of a size which would have led to different conclusions. This strongly suggests both that the analysis sample is reasonably representative of the total cohort, and that any biases do not affect statistical relationships. We cannot discount the possibility of systematic bias throughout, such as the under-representation of ethnic minorities, but it seems extremely unlikely that any such groups would be large enough to change the overall relationships.

RESULTS

Summary of bivariate results

With few exceptions there was a regular correlation between truancy and the outcomes identified below, with truants at 16 only in an intermediate position between truants at 11 and 16 and non-truants.

Regarding occupational outcomes up to age 23, truancy was associated with an unstable job history: a shorter mean length of jobs, and a higher total number of jobs. There was not on the whole less job satisfaction. Truants were more often unemployed, with the proportion of economically active time unemployed and the mean length of all unemployment spells also being longer. For those who were in work, there was found to be little difference in pay between truants and non-truants, with the exception of female truants who earned somewhat less in their current job, though not their first job. However, truancy was associated with a lower family income, a measure which takes into account family size and structure. Tables and a fuller discussion of these findings are available in Hibbett (1987).

Multivariate analysis

These findings, interesting though they are, cannot establish any causal influence of truancy on occupational outcomes. Although there were found to be certain differences between truants and non-truants, these could be due to other factors, like social background and educational ability. The multivariate analyses described below attempt to take such factors into account and thus come closer to identifying the particular influence of truancy, and its nature.

Two kinds of analyses were used: analyses of variance where the outcome variable was continuous (for example, pay in current job); and loglinear analysis where the outcome variable was categorical (for example, economic status).

Irrespective of the kind of analysis and the particular outcome, the structure and sequence of analyses was the same for all outcome variables, and designed to test the possible explanations of the initial association between truancy and outcome as discussed in the introduction. In all cases the contrast of interest was between the truants and the non-truants, according to the teachers' ratings at 16. Firstly, the simple contrast between the two groups was produced; the second analysis included social background factors (social class of father's occupation, region of residence, family size) as independent variables and adjusted the truancy contrast for the relationship with those factors; the third stage additionally incorporated and adjusted for measure of ability and attainment at age 11 (general ability, reading comprehension, mathematics test); the fourth stage also included the subject's attendance rate at 15; and the fifth and final stage also included a measure of public examination attainment by the time of leaving school.

Thus at each stage it was possible to identify the particular contribution to the relationship between truancy and the outcome of the variable, or the set of variables, added in at that stage. Because of the differing examination systems between Scotland and the rest of Great Britain, it was not possible to devise a common scale of examination attainment for the final stage of the analysis. As the Scottish sample was too small to provide reliable results in the loglinear analyses, they have been excluded here. Findings for the Scottish respondents from the analyses of continuous variables can be found in Hibbett and Fogelman (1988). For all the above, separate analyses were carried out for men and women.

Tables are presented for main outcomes; others which are mentioned are available in Hibbett and Fogelman (1988). The tables summarise all the stages of analysis described above for any one outcome.

Occupational outcomes

Table 2 compares the current (i.e., at age 23) economic status of the truants and non-truants. This and subsequent tables based on loglinear analysis report the results as relative odds. For each outcome variable, one category was selected as reference category against which the relative odds of the other categories are calculated. 95 per cent confidence intervals are presented for each stage of the subsequent analyses in order to indicate statistical

TABLE 2

ECONOMIC STATUS (LOGLINEAR ANALYSIS). ODDS RELATIVE TO WORK: TRUANTS AT 16 V. NON-TRUANTS

	Adjusted for:	Background	AND Educ. Ability at 11	AND Attendance at 15	AND Qualifications at 16
MEN (2806)					
Full-time Education	0.45 (0.22/0.90)	0.55 (0.27/1.12)	0.57 (0.26/1.26)	0.58 (0.26/1.31)	0.59 (0.26/1.33)
Unemployed	2.37 (1.84/3.14)	2.03 (1.54/2.68)	2.04 (1.54/2.71)	2.05 (1.54/2.72)	2.04 (1.53/2.71)
Out of Labour Force	1.43 (0.74/2.77)	1.56 (0.79/3.11)	1.60 (0.76/3.39)	1.62 (0.76/3.45)	1.60 (0.75/3.41)
WOMEN (2843)					
Full-time Education	0.46 (0.12/1.76)	0 54 (0.12/2.51)	0.52 (0.09/2.93)	0.53 (0.10/2.66)	0.61 (0.13/2.81)
Unemployed	2.76 (1.90/4.0)	2.32 (1.57/3.41)	2.42 (1.63/3.57)	2.42 (1.64/3.59)	2.41 (1.63/3.56)
Out of Labour Force	3.29 (2.62/4.14)	2.59 (2.04/3.28)	2.67 (2.08/3.43)	2.68 (2.09/3.44)	2.56 (1.99/3.31)

(95 per cent confidence intervals in brackets)

significance. For all loglinear analyses, interactions between the truancy measure and other independent variables were tested. However, very few reached statistical significance and none was consistent and readily interpretable. Thus main effects only are fitted in the final models.

In Table 2, the first column shows, for men, that before any adjustments for other factors truants compared to non-truants are: 0.45 times as likely to be in full-time education rather than in work; 2.4 times as likely to be unemployed rather than in work; and 1.4 times as likely to be out of the labour force rather than in work. Allowing for social background reduces the contrasts in full-time education and unemployment, though not substantially. The confidence intervals associated with both full-time education and being out of the labour force indicate that it would be unsafe to conclude that the truants are significantly different in these respects, but the contrasts in the odds of being unemployed are highly significant. These differences appear to be unaffected by allowing for further factors, suggesting that the increased odds of unemployment for truants are not explained by their poor attendance or their poor examination qualifications, once social background has been taken into account.

The patterns for women are very similar, except that the truants' relative odds of being out of the labour force are considerably greater. The confidence intervals associated with this are smaller than for men, as this is of course more common for women of this stage, including as it does those who are not working and not seeking work because they are caring for children. It is possible that this group includes some who report that they are not seeking work, but would do so if they thought that suitable work was available. This may therefore reflect some additional "hidden" unemployment.

TABLE 3

SOCIAL CLASS OF CURRENT JOB (LOGLINEAR ANALYSIS). ODDS RELATIVE TO 'UNSKILLED MANUAL' TRUANTS AT 16 V. NON-TRUANTS

	Adjusted for:				
		Background	AND Educ. Ability at 11	AND Attendance at 15	AND Qualifications at 16
MEN (2666)					
I + II	0.12 (0.08/0.20)	0.18 (0.11/0.29)	0.18 (0.11/0.29)	0.17 (0.10/0.29)	0.17 (0.10/0.29)
III + IV Non-Manual	0.11 (0.07/0.18)	0.15 (0.09/0 25)	0.15 (0.09/0.26)	0.15 (0.09/0.26)	0.16 (0.09/0.26)
III + IV Manual	0.59 (0.41/0.85)	0.66 (0.45/0.96)	0.65 (0.45/0.96)	0.64 (0.43/0.94)	0.64 (0.43/0.95)
WOMEN (2776)					
I + II	0.07 (0.04/0.13)	0.11 (0.06/0.21)	0.13 (0.06/0.28)	0.13 (0.06/0.28)	0.12 (0.05/0.26)
III + IV Non-Manual	0.12 (0.07/0 22)	0.16 (0.09/0.30)	0.16 (0.08/0.32)	0.16 (0.08/0.31)	0.16 (0.08/0.32)
III + IV Manual	0.51 (0.28/0.90)	0.53 (0.29/0.96)	0 55 (0.28/1.06)	0 54 (0.28/1.05)	0.56 (0.29/1.10)

(95 per cent confidence intervals in brackets)

Table 3 examines the social class of current job (or most recent job for those not currently working). Broadly, the results are very similar for each sex (and also, not presented here, for the social class of first job on completing education). Relative to being in unskilled manual occupations truants are at 23 about one tenth as likely to be in professional, managerial or intermediate non-manual occupations (classes I and II), a fifth to a quarter as likely to be in skilled or semi-skilled non manual occupations (III and IV Non-manual), and a half to two thirds as likely to be in skilled or semi-skilled manual occupations (III and IV Manual). As for economic status, the contrasts are slightly reduced by allowing for social background, but then barely affected by the introduction of further factors into the analysis.

There is a slight suggestion, at least with regard to the distinction between skilled or semi-skilled as against unskilled jobs — which are the main destinations of the truants — of a widening of the gap in the period between first and current jobs. For men, relative odds for being in semi-skilled or skilled rather than unskilled manual work were 0.8 for the first and 0.6 for the current job; for women the respective odds were 0.9 compared to 0.6.

TABLE 4

NUMBER OF FULL- AND PART-TIME JOBS (ANALYSIS OF VARIANCE). DEVIATIONS FROM OVERALL MEAN

	Adjusted for:	Background	AND Educ. Ability at 11	AND Attendance at 15	AND Qualifications at 16+
MEN (2815)					
Truant at 16 (525)	1.09	0.99	0.87	0.74	0.57
Not Truant (2290)	-0.25	-0 23	-0.20	-0.17	-0.13
Mean Difference	1.34***	1.22***	1.07***	0.91***	0.70**
WOMEN (2848)					
Truant at 16 (398)	0.51	0.53	0.49	0.42	0.33
Not Truant (2450)	-0.08	-0.09	-0.08	-0.07	-0.05
Mean Difference	0.59***	0.62***	0.57***	0.49***	0.38***

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

Table 4 compares the number of jobs held by truants and non-truants. Interpretation of this and subsequent tables based on analysis of variance is relatively straightforward, as the figures represent differences on the scale of the outcome measure. For example, in Table 4, the mean differences refer to number of jobs. For each stage of the analysis, the probability level is given of the F-value associated with the difference between the truants and non-truants. As for the loglinear analyses, interactions between the truancy variable and other independent variables in the analysis of variance were tested, but found to be generally insignificant and inconsistent. Truants have, by the age of 23, held more jobs than have non-truants. However, this contrast is reduced at each stage of the analysis; that is, each of the other factors contributes to the explanation of the difference. Nevertheless, the final adjusted difference of 0.7 jobs for men and 0.4 jobs for women remains statistically significant.

Of course differences in the number of jobs held could simply reflect the fact that truants will have left school earlier and been in the labour market longer, rather than necessarily any tendency to change jobs more frequently. That truants have been in employment longer is confirmed, at least for men. Table 5 shows that, on average, the former truants have been in employment some 15 months longer by 23. However this difference is reduced to only five months when social background and attainment at 11 are taken into account, and further reduced by the introduction of attendance rate and examination performance, to where it is no longer statistically significant.

For women, the pattern is very different. The raw mean difference in months of time in employment between truants and non-truants is very small (with truants having had slightly less time in employment), but increases when other factors are allowed for. It is extremely likely that this is the result of departure from the labour force as seen in Table 2. Female non-truants' time in the labour force is truncated because they stay in education, and truants' because they leave employment when they marry and have children. These two effects are complementary and more or less offset each other in terms of the raw difference, but once the truants are compared with others similar in their social background, attainment at 11 and examination attainment (attendance contributing relatively little), they have spent considerably less time in employment.

Looking at the average duration of all jobs, truants have on average held their jobs for a shorter time (Table 6). Furthermore, the contrast with non-truants is generally increased as account is taken of further factors. The final difference in average job length, of nine months for men and 11 months for women, is quite dramatic given the relatively short time all cohort members can have been in employment by the age of 23.

TABLE 5

TOTAL MONTHS IN FULL- AND PART-TIME JOBS (ANALYSIS OF VARIANCE). DEVIATIONS FROM OVERALL MEAN

	Adjusted for:	Background	AND Educ. Ability at 11	AND Attendance at 15	AND Qualifications at 16+
MEN (2514)					
Truant at 16 (397)	13.01	9.57	4.28	3.62	-2.18
Not Truant (2117)	-2.44	-1.79	-0.80	-0.68	0.41
Mean Difference	15.45***	11.36***	5.08**	4.30*	2.59
WOMEN (2620)					
Truant at 16 (349)	-1.25	-3.30	-6.51	-6.77	-10.10
Not Truant (2271)	0.19	0.51	1.00	1.04	1.55
Mean Difference	1.43	3.81	7.51	7.81	11.65**

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

Men who had played truant from school at 16 have, by 23, experienced an average of 4.2 months more unemployment (table not presented). Each of the factors introduced into the analysis explains a portion of this difference, to the point where it is reduced to 2.3 months but still highly statistically significant. The pattern for women is very similar, but with a slightly smaller difference.

TABLE 6

MEAN LENGTH OF ALL FULL- AND PART-TIME JOBS (MONTHS) (ANALYSIS OF VARIANCE), DEVIATIONS FROM OVERALL MEAN

	Adjusted for:	Background	AND Educ. Ability at 11	AND Attendance at 15	AND Qualifications at 16+
MEN (2666)					
Truant at 16 (499)	-3.5	-4.0	-5.7	-4.6	-7.3
Not Truant (2167)	0.8	0.9	1.3	1.1	1.7
Mean Difference	4.3***	4.9	7.0*	5.7	9.0**
WOMEN (2689)					
Truant at 16 (368)	-6.0	-7.2	-8.9	-8.1	-9.7
Not Truant (2321)	0.9	1.1	1.4	1.3	1.5
Mean Difference	6.9***	8.3**	10.3***	9.4***	11.2***

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$

Of course, as a measure of amount of unemployment, total length of all spells will again be confounded with the length of time in the labour market. A better measure which takes this into account is time unemployed as a proportion of total time in the labour market (Table 7). This variable deviated from normality for both sexes and from the assumption of homogeneity of variance only for women. Although the F-test in the analysis of variance is robust with respect to assumptions of normality and homogeneity of variance (Gaito, 1973; Winer, 1975), we repeated the analysis with a transformed variable, with results being similar to those reported in Table 7. Before the introduction of other factors, truants have spent just less than 2 per cent more of their economically active time unemployed. The effect on this of allowing for the other factors is not very consistent or very great, but comparing with others of similar social background does reduce the contrast enough to make it no longer significant at the 5 per cent level. However, the contrast is increased again when adjusted for differences in examination attainment. Far from explaining the greater unemployment of the truants, their disadvantage is greater than others of similar attainment.

TABLE 7

TOTAL TIME UNEMPLOYED AS A PERCENTAGE OF ECONOMICALLY ACTIVE TIME (ANALYSIS OF VARIANCE). DEVIATIONS FROM OVERALL MEAN

	Adjusted for:				
		Background	AND Educ Ability at 11	AND Attendance at 15	AND Qualifications at 16+
MEN (2767)					
Truant at 16 (518)	1.6	1.4	1.9	1.2	2.5
Not Truant (2249)	-0.4	-0.3	-0.4	-0.3	-0.6
Mean Difference	2.0**	1.7	2.3	1.5	3.1*
WOMEN (2806)					
Truant at 16 (393)	4.1	3.3	3.4	3.0	3.2
Not Truant (2413)	-0.7	-0.5	-0.6	-0.5	-0.5
Mean Difference	4.8***	3.8**	4.0**	3.5**	3.7**

* P < 0.05; ** P < 0.01; *** P < 0.001

For women, the contrast in the proportion of time unemployed is greater, almost 5 per cent before allowing for other factors. It is reduced slightly by allowing for background and then barely affected by the subsequent stages of analysis.

However, these differences amount to relatively small proportions of the total economically active time. Looking at the mean length of all unemployment spells provides further confirmation that the unemployment histories of the truants are not markedly worse. The raw average difference amounts to 1.5 months for men and 2.3 months for women, but this is reduced to a non-significant level when allowance is made for social background and ability (background alone for the women).

Earnings and income

In examining financial outcomes we first compared weekly earnings in the first job of truants and non-truants. Of course, a straight comparison in this respect would be extremely misleading because of the varied timing of entry into the labour market and the effects of inflation. These analyses were therefore limited to those who took their first job in 1974, when they were 16. Average differences in earnings at that time were not very large, with the truants in fact earning slightly more. For men the raw difference is not statistically significant, and is further reduced as, in particular, attendance rate and qualifications are allowed for. For women the initial differences are larger (again in favour of the former truants), but still only border on significance once social background is taken into account.

Results are very similar when the current earnings of those in employment are examined (Table 8). Among men, truants earn slightly more on average and the difference is increased when other factors are taken into account, but it is not statistically significant throughout these analyses. Among women, the raw difference in average earnings is quite large (just over £7 per week), and truants earn less. However, this is reduced to a few pence when social background and ability are taken into account, and the contrast is reversed, but still not statistically significant, when allowance is made for attendance and examination attainment.

These results cannot be taken to mean that those who had played truant are no worse off than others. As explained, they are limited to those currently working. The unemployed and those out of the labour force are therefore excluded, both accounting for disproportionate numbers of truants, as we saw in Table 2. Furthermore, the meaning of a particular income will differ according to other circumstances. The same amount would feel quite different to, say, someone who is single and still living with their parents, and someone who is married, has children and is paying for their own home. The latter was found to be more likely to be the situation at 23 of those who had played truant at school.

TABLE 8

CURRENT JOB PAY PER WEEK (£s) (ANALYSIS OF VARIANCE). DEVIATIONS FROM OVERALL MEAN

	Adjusted for.				
		Background	AND Educ. Ability at 11	AND Attendance at 15	AND Qualifications at 16+
MEN (2285)					
Truant at 16 (394)	1.4	1.7	2.6	2.6	2.1
Not Truant (1891)	-0.3	-0.4	-0.6	-0.5	-0.4
Mean Difference	1.7	2.1	3.2	3.1	2.5
WOMEN (1833)					
Truant at 16 (166)	-6.4	-4.3	-0.2	0.8	1.8
Not Truant (1667)	0.6	0 4	0.0	-0.1	-0.2
Mean Difference	7.0***	4.7***	0.2	0.9	2.0

* P < 0.05; ** P < 0.01; *** P < 0.001

For these reasons, a better measure of current financial circumstances is provided by "equivalent net family income" as examined in Table 9. This measure takes into account income to the family (i.e., the unit of which the cohort member is head or joint head, thus not including parents etc. if not yet established as an independent household) from all sources. This measure is then weighted according to the structure and size of the family, using the weightings (e.g., for children of different ages) incorporated in calculations of Supplementary Benefit (Shepherd, 1984).

TABLE 9

EQUIVALENT NET FAMILY INCOME PER WEEK (£s) (ANALYSIS OF VARIANCE). DEVIATIONS FROM OVERALL MEAN

	Adjusted for:				
		Background	AND Educ. Ability at 11	AND Attendance at 15	AND Qualifications at 16+
MEN (2658)					
Truant at 16 (498)	-12.0	-8.0	-6.0	-3.5	-4.8
Not Truant (2160)	2.8	1.8	1.4	0.8	1.1
Mean Difference	14.8***	9.8	7.4	4.3	5.9
WOMEN (2751)					
Truant at 16 (386)	-23.3	-16.5	-9.7	-7.8	-5.8
Not Truant (2365)	3.8	2.7	1.6	1.3	1.0
Mean Difference	27 1***	19.2***	11.3**	9.1*	6.8*

* P < 0.05; ** P < 0.01, *** P < 0.001

On this measure truants are considerably less well off, almost £15 per week on average for men and £27 per week for women. For both sexes, allowing for other factors reduces this considerably. For men, adjusting the comparison for social background alone brings it down to less than £10, which is no longer significant at the 5 per cent level. For women, the contrast remains greater. It is for example still over £11 on average when comparing with those of similar background and tested ability. Allowing additionally for attendance and examination attainment reduces the difference to £6.75 per week and the borderline of statistical significance.

DISCUSSION

Before discussing the findings, it is important to note that their interpretation must be in the context of this particular cohort, when it was studied, and their age at that time.

NCDS members could first leave school in 1974 (they were in fact the first year group to be affected by the raising of the school leaving age to 16). Their first entry into the labour market could be from that year through to 1981 — a small number had still not taken up employment by the time of the 23-year interview. Thus, although unemployment levels were steadily rising during this time, they will not have experienced the dramatic rise in unemployment, with its disproportionate effect on school-leavers, encountered in the 1980s. Equally, their entry into employment generally pre-dated those schemes which were introduced to assist young school-leavers, and also school-based initiatives such as TVEI. On the other hand, other job-related training opportunities, most notably apprenticeships in manufacturing industries, have diminished in the intervening years.

Thus, the employment opportunities for this cohort, at least in their first few years after completing their education, were certainly greater than for more recent year groups. If, as the preliminary results suggested, those who had truanted from school were experiencing difficulties in the labour market in the late 1970s, we might expect these to be increased by subsequent economic difficulties. It seems likely that any contrasts we may identify for the 1958 cohort will be, if anything, greater for younger age groups.

The second general point relevant to all our findings concerns the age at which our most recent data were collected, that is 23. This is clearly relatively early in the adult lives of the cohort (although it is also likely to be at a relatively early age that the impact of school truancy is most direct and less affected by other life events). By this age most will have entered the labour market, but many only relatively recently. In particular those who continued into higher or further education, though working, may not yet have settled into the careers which will typify their working lives. Their current earnings will be relatively low compared to their lifetime potential, whereas manual workers' earnings will, in real terms, vary relatively little as they become older.

Most importantly, the rate and sequence of progress through the transitions of early adulthood are not random phenomena. There are substantial differences according to class of origin and according to sex (Kiernan, 1986). Young people from working class backgrounds tend to leave school and enter the labour market at an earlier age than those from middle class backgrounds.

In basing this study on information which covers the period up to the age of 23, we are describing an age when progress through the life cycle varies substantially, and this variation relates to background and sex. Since it is known that truants differ in similar ways, these general patterns will help to explain some of the contrasts which we find between truants and non-truants, how this varies between the sexes, and how accounting for background variables affects the results. These points also reinforce the case for the kind of multivariate analysis which has been carried out.

Our findings have clearly demonstrated that young people who had been identified by their teachers as truanting from school at the age of 16 could be distinguished from others in terms of their occupational characteristics in the years after leaving school and up to the age of 23. They have lower status occupations, less stable career patterns and more unemployment. Among those who are working, their incomes from employment are not lower (indeed among men they are a little higher), but they are considerably less well off once their family situation is taken into account. Furthermore, their apparently high earnings are probably somewhat artificial as they are mainly in manual jobs and at or near the peak of their earning power, while many of their peers, particularly those who have not long finished their education, will substantially increase their earnings as they grow older.

Although, because of small numbers, it was not possible to confirm this through multivariate analysis, the simpler relationships in the preliminary analyses suggest that such differences are even more marked for those who had additionally been identified by their teachers as truanting towards the end of their primary schooling.

However, it is known that truants are drawn disproportionately from children of less advantaged backgrounds and of lower ability (Farrington, 1980; Fogelman and Richardson, 1974). It would be possible that this alone might explain some or all of the above; that truants' early occupational experiences are no different when they are compared with other young people of similar backgrounds and ability. The analyses described above were designed in part to assess whether this was the case. In fact, they suggest that these factors do explain some part of the differences in the early adult situations of truants and non-truants, but generally only a small part. The two areas where it does appear that the raw differences are largely accounted for in this way are unemployment history (as distinct from current unemployment at 23) and income. For example, on what is probably the most useful unemployment history measure — time unemployed as a proportion of economically active time — initial differences were fairly small (truants having been unemployed for about 2 per cent more of their economically active lives), and the small reduction when social background was allowed for was sufficient to make this no longer statistically significant. Raw differences in family income showed truants much less well off at 23, but, for men only, this was not at a statistically significant level once social background was taken into account. Female truants do, however, appear to have considerably lower family incomes than others of similar background and tested ability.

For most occupational outcome variables, social background contributed not at all to explaining the contrasts between truants and non-truants, or only a small amount, which still left a substantial and significant difference between the two groups.

The contrasts in occupational experience are substantially in agreement with those found by Cherry (1976) for the cohort born in 1946, despite the changes which there would have been in the labour market in the intervening 12 years. As for the 1958 cohort, she found that former truants changed jobs more frequently, but that their average earnings were not lower. In this last respect both studies contradict the American findings of Robins and Ratcliffe (1980).

Some previous research in this area has not made the distinction between truancy and absenteeism (see for example Robins and Ratcliffe). It has also been suggested that the distinction is not important in the context of subsequent effects, as it is missing school, irrespective of the reasons, that leads to subsequent problems.

If this were the case, then we should expect to find that, after adjusting for differences in background and ability, further adjustments for attendance rate would explain any remaining differences between truants and non-truants. In other words, we would be testing whether truants differ in their outcomes from others whose attendance was equally poor, but who were not labelled truants. Our analyses offer little support for such a hypothesis. The effect on the contrast between truants and non-truants (of similar social background and early attainment) of allowing for attendance rate in the autumn term of the final compulsory year is not consistent and rarely more than trivial. In no case did it prove to explain a previously significant difference. Thus, it does seem that truants are distinct from other poor attenders in terms of their occupational outcomes in early adulthood.

Absence from school, particularly in the final years, is associated with depressed attainment (Fogelman, 1978). Truants do leave school with fewer qualifications. It does appear likely that subsequent difficulties in the labour market could result from this, as was claimed by Gray *et al.* (1980). The NCDS findings do not support this. Allowing for end-of-school qualifications does have some effect on labour market outcomes associated with truancy, but not substantial nor sufficient to explain the differences found. Indeed for two of the main indicators of employment experience, average job length and proportion of economically active time unemployed, the introduction of qualifications into the analysis increased the contrast between truants and the rest. Former truants appear to have more unemployment and job instability than others of similar social background and ability who leave school with similar qualifications (or, more commonly, similarly unqualified).

CONCLUSIONS

The underlying purpose of these analyses has been to attempt to cast light on competing explanations for truancy from school.

The results do not seem compatible with those explanations which suggest that truants have simply outgrown school and are ready for the world of work. The evidence presented here does not suggest that they have no further problems once school is behind them and that they progress smoothly and happily through employment. On the contrary, they experience more unemployment, more job changes and less job satisfaction. At the age of 23 they are more than twice as likely to be unemployed, and the average duration of each job held is nine months less for men and 11 months less for women. These last figures are particularly striking when we bear in mind that they left school and entered employment at most seven years previously.

These outcomes seem more compatible with the explanations proffered by those who characterise truancy as evidence of more comprehensive psychological problems (Tyerman, 1968). Perhaps the "reluctant adolescent" (Haigh, 1976) has grown up to be a reluctant adult.

This is not to suggest that the problem is located solely in the individual. The school system may be such, through internal organisation, competitive examinations, etc., that some individuals will inevitably feel undervalued and alienated. This study has shown that there are continuities in this respect with adult life, and that the difficulties experienced by those who withdraw from school are greater than would be predicted by their school failure alone. Research on school effectiveness is beginning to identify ways in which the organisation and content of schooling may be related to disaffection and truancy levels (Reynolds and Murgatroyd, 1977; Rutter *et al.*, 1979; Reynolds, 1985; Mortimore *et al.*, 1988). The Hargreaves Committee (Inner London Education Authority (ILEA), 1984) produced a number of detailed recommendations for improving school attendance, but also acknowledged that only "part of the cause of truancy lies within the schools' control". It is not realistic to expect that schools can structure themselves and their curricula to ensure a successful and fulfilling experience for every one of their pupils.

This is not to imply that schools and their supporting services should ignore truancy. On the contrary, these findings do show that for the individual it is often a sign of more general and longer-term difficulties, and a predictor of employment problems of a more severe kind than will be experienced by others who share the disadvantaged background and low attainment which typify the truant.

ACKNOWLEDGEMENTS.—We are grateful to the ESRC who funded this study, and to our colleagues in the Social Statistics Research Unit, City University, for their advice and support.

Correspondence and requests for reprints should be addressed to Dr Angelika Hibbett, Employment Department, Steel House, Tothill Street, London SW1H 9NF

REFERENCES

Butler, N. R., and Bonham, D. G. (1963). *Perinatal Mortality*. Edinburgh: Livingstone.

Cherry, N. (1976). Persistent job changing — Is it a problem? *J. occup. Psychol.*, 49, 203-221.

Davie, R., Butler, N. R., and Goldstein, H. (1972). *From Birth to Seven*. London: Longman.

Farrington, D. (1980). Truancy, delinquency, the home and the school. In Hersov, L., and Berg, I. (Eds.), *Out of School* Chichester: John Wiley and Sons.

Fogelman, K. (1978). School attendance, attainment and behaviour. *Br. J. educ. Psychol.*, 48, 148-158.

Fogelman, K. (Ed.) (1983). *Growing Up in Great Britain*. London: Macmillan.

Fogelman, K., and Richardson, K. (1974). School attendance: some results from the National Child Development Study. In Turner, B. (Ed.), *Truancy*. London: Ward Lock Educational.

Gaito, J. (1973). *Introduction to Analysis of Variance Procedures*. New York: MSS Information Corporation.

Goldstein, H. (1983). A study of response rates of 16-year-olds in the National Child Development Study. In Fogelman, K. (Ed.), *Growing Up in Great Britain*. London: Macmillan.

Gray, G., Smith, A., and Rutter, M. (1980). School attendance and the first year of employment. In Hersov, L. and Berg, I. (Eds.), *Out of School*. Chichester: John Wiley and Sons.

Haigh, G (1976). *The Reluctant Adolescent*. Aldershot: Temple Smith.

Hibbett, A. (1987). Early Adult Outcomes of Truancy. NCDS User Support Group Working Paper 24, Social Statistics Research Unit, City University.

Hibbett, A , and Fogelman, K. (1988). Early Adult Outcomes of Truancy II: The effects of truancy after allowing for other factors. NCDS User Support Group Working Paper 30, Social Statistics Research Unit, City University.

Inner London Education Authority (1984). *Improving Secondary Schools*. London: ILEA.

Iyer, R. (1984). Analysis of Response. NCDS4 Working Paper 25, NCDS User Support Group, Social Statistics Research Unit, City University.

Kiernan, K. (1986). Transitions in Young Adulthood. NCDS User Support Group Working Paper 16, Social Statistics Research Unit, City University

Mortimore, P. *et al.*, (1988). *School Matters*. Wells: Open Books.

Reynolds, D. (1985). *Studying School Effectiveness*. London: Falmer Press.

Reynolds, D., and Murgatroyd, S. (1977). The sociology of schooling and the absent pupil. In Carroll, H. C. M. (Ed.), *Absenteeism in South Wales*. Swansea: Faculty of Education, University of Swansea.

Robins, L. N., and Ratcliffe, K. S. (1980). The long-term outcome of truancy. In Hersov, L. and Berg, I., (Eds.). *Out of School*. Chichester: John Wiley and Sons.

Rutter, M., *et al.* (1979). *Fifteen Thousand Hours*. Wells: Open Books.

Rutter, M., Tizard, J., and Whitmore, K. (1970). *Education, Health and Behaviour*. London: Longman.

Shepherd, P. (1984). Earnings, income and other aspects of the financial circumstances of the NCDS Cohort at 23. NCDS4 Working Paper 19, NCDS User Support Group, Social Statistics Research Unit, City University.

Stott, D. H. (1969). *The Social Adjustment of Children*. London: University of London Press.

Tyerman, M. J. (1968). *Truancy*. London: University of London Press.

Winer, B. J. (1975). *Statistical Principles in Experimental Design*. London: McGraw Hill.

(*Manuscript received 3rd August*, 1988)

Br. J. educ. Psychol., 60, 37-51, 1990

PSYCHOLOGICAL AND BIOGRAPHICAL DIFFERENCES BETWEEN SECONDARY SCHOOL TEACHERS EXPERIENCING HIGH AND LOW LEVELS OF BURNOUT

BY C. MARK B. PIERCE AND GEOFFREY N. MOLLOY
(Monash University, Victoria, Australia)

SUMMARY. A total of 750 teachers from 16 government and non-government schools from areas of contrasted socio-economic status (SES) responded to a questionnaire designed to investigate associations between selected aspects of burnout among teachers working in secondary schools in Victoria, Australia. By comparing high and low burnout groups on biographic, psychological and work pattern variables, differences between teachers experiencing high and low levels of burnout were identified. Multiple regression analyses assessed the relative importance of these variables in accounting for the variance in each of the three burnout subscales. School type was related to perceptions of stress and burnout. Higher levels of burnout were associated with poorer physical health, higher rates of absenteeism, lower self-confidence and more frequent use of regressive coping strategies. Teachers classified as experiencing high levels of burnout attributed most of the stress in their lives to teaching and reported low levels of career commitment and satisfaction. Further, teachers who recorded high levels of burnout were characterised by lower levels of the personality disposition of hardiness, lower levels of social support, higher levels of role stress and more custodial pupil control ideologies than their low-burnout counterparts. Psychological variables were found to be more significant predictors of burnout than biographical variables.

INTRODUCTION

IN Selye's (1976) scheme, "stress" can have both positive and negative effects. When people experience moderate amounts of stress they feel euphoric and stimulated, and Selye, with his penchant for Greek, described this positive experiential state as "eustress". In contrast, a person experiencing too much (or too little) stress is more likely to feel dysphoric, tense and fatigued, and this has been referred to as negative stress or "dystress". In the literature, the terms, dystress and stress are synonymous.

It is well documented that teaching is regarded as a stressful occupation and recurring stressors in the teaching environment have been reported by numerous writers (e.g., Coates and Thoresen, 1976; Kyriacou and Sutcliffe, 1977, 1978b, 1979a, 1979b; Cichon and Koff, 1978; Cooper and Marshall, 1980; Dunham, 1980; Otto, 1980, 1982, 1986; Fisher, 1984; Laughlin, 1984a, 1984b). These include high work loads and time pressures, problems with students, staff conflicts, high noise levels, overcrowded classrooms and staffrooms, lack of recognition for work done and encountering negative community attitudes towards the teaching profession.

One way in which teachers react to these stressors is manifested in a syndrome called "burnout". The term burnout has been used to describe extreme stress responses which are specifically linked to the emotional strain of dealing extensively with other people (Maslach, 1976). There appear to be at least three aspects of teacher burnout. The first is the development of increased feelings of emotional exhaustion and fatigue. Second, is the tendency for teachers to develop negative, cynical attitudes toward their students. The third aspect of burnout is the tendency to evaluate oneself negatively, resulting in feelings of lack of personal accomplishment.

Many factors are likely to contribute to the development of burnout. These can be grouped into three broad categories, namely, personal, interpersonal and organisational factors. The question being addressed at the individual level is, "What role do individual characteristics of the teacher play in burnout?" Individual factors which may contribute to burnout include certain personality traits and background characteristics. For the individual teacher, feelings of burnout may result in both psychological and physical dysfunction and this may be reflected in feelings of exhaustion, depression and irritability, increased use of alcohol and medication, and increased susceptibility to illness.

At the interpersonal level the question of interest is, "What role do other people (students, colleagues, parents, community members) play in burnout?" Since the teacher is surrounded by other people, this question is directed at looking at the impact on teachers of the social environment. The nature of teachers' involvement with students, colleagues, parents and members of the community are posited to affect self-perceptions of burnout. If teachers experience tensions and difficulties associated with this involvement, then this may be reflected in their behaviour towards others.

Third, at the organisational level, the question can taken two forms, specifically, "What role does the school setting play in burnout?" and, "What role does the employer play in burnout?" Ways in which the school setting contributes to burnout include a lack of clarity in the definition of teachers' duties, vague rules and regulations, large classes, poor availability of resources, cramped working conditions and lack of privacy. The effects of burnout at the school level may be reflected in staff turnover, absenteeism and teacher dissatisfaction. Employer contributions to burnout include poor communication and implementing decisions which affect teachers without having consulted them, inflexible policies on staffing and leave, inadequate teacher welfare services, and a general lack of formal support systems for teachers. The effects of burnout at the employer level may include high resignation rates with a consequent loss of trained and experienced teachers, increased numbers of "leave without pay" applications and workers' compensation claims.

Factors which are likely to be associated with burnout include individual, interpersonal and organisational variables. Only a partial picture of the burnout process would be obtained if individual variables were studied in isolation without consideration of the social and organisational context in which teachers work. Some of the variables likely to be associated with levels of burnout have been suggested in the extant literature. Recent American and British studies have found that role conflict and role ambiguity contributed to feelings of burnout in teachers (Schwab and Iwanicki, 1982; Pierson, 1984; Capel, 1987). However, no studies have examined these relationships with a sample of Australian teachers. Laughlin (1984a) reported that social support reduced perceptions of stressors among Australian teachers. Although these results suggest that social support may attenuate feelings of burnout, no studies were located in the literature which have directly examined the relationship between social support and burnout.

In the United States, Kobasa's research (1979, 1982b) on the relationship between stressful life events and illness onset has identified the personality disposition of hardiness as being a stress resistance resource. The hardy personality comprises the three dimensions of commitment, challenge and control. The commitment dimension refers to a generalised sense of purpose and meaningfulness about life that minimises the perceived threat of any given stressor. Committed persons become actively involved in the task at hand, do not give up in demanding times and seek assistance when under pressure. Control refers to the belief in one's ability to have an influence on situations that arise rather than feeling helpless when confronted with unpleasant events. The challenge dimension refers to the expectation that it is normal for life to change. Change, for people scoring high on this factor, is perceived as providing an opportunity for development and growth rather than a burden and a threat to security.

When Kobasa considered people who had experienced comparable levels of stressful life events, it was found that those who were high in hardiness reported less illness than those who were low in hardiness. Similarly, hardiness may attenuate the effects of job-related stress and contribute to reduced levels of teacher burnout. Teachers with a hardy personality disposition may experience lower levels of burnout as a consequence of their positive cognitive appraisal of stressors and the use of decisive coping actions to deal with them. However, the relationship between hardiness and burnout has not been investigated.

Kobasa (1982a) investigated stress resistance resources among a sample of lawyers and found a positive relationship between strain symptoms and the use of regressive coping strategies. Regressive coping strategies are identified as behaviours which attempt to deny, minimise, or avoid stressful situations and are regarded as being ineffective ways of coping with stressors. Since a relationship between strain symptoms and regressive coping strategies has been found, it is likely that higher levels of burnout will be related to more frequent use of regressive coping strategies.

Teachers' pupil control ideologies have been found to be related to levels of anxiety (Docking and Docking, 1984; Docking, 1985). Teachers with a custodial orientation were found to be more anxious than those with a less custodial orientation. Again, while this finding suggests that pupil control ideology may be related to burnout, this relationship has not been examined.

Research is often focused on the individual teacher without consideration of the environment in which the teacher works. According to the classification of organisations outlined by Carlson (1964), non-government schools are characterised by voluntary attendance and organisational selectivity of students whereas government schools are characterised by mandatory provision of the service and involuntary student attendance. Moreover, within these two categories, the environments of schools may vary according to socio-economic status (SES) level. Hence, teachers may work in different environments and under distinctive conditions. Consequently, it is quite conceivable that the type of school in which teachers work may be related to levels of burnout.

The important influence of the SES level of the area in which teachers work on their stress levels is suggested by the results of a study carried out by Pratt (1978) among a sample of primary school teachers in England. The results of this study indicated that teachers working in schools in low SES areas recorded higher levels of stress and lower levels of psychological well-being as measured by the General Health Questionnaire (GHQ) than teachers working in schools in high SES areas.

The finding that stress tended to increase as the age of the children taught increased implies that this effect may be greater in secondary schools. Fisher (1983, 1984) found that classroom teachers reported lower levels of personal accomplishment than teachers with additional responsibilities. This indicates a relationship between a teacher's position in the school and feelings of burnout. Further, since teachers with fractional time appointments are exposed to the stressors of teaching for shorter time periods it is likely that they would experience lower levels of burnout than full-time teachers.

A negative relationship between burnout and self-concept has been reported among prison officers (Schlachter, 1981) and among student primary teachers (Blough, 1981). Considering self-confidence as a specific aspect of self-concept, it would be expected that there would be a negative association between burnout and self-confidence among secondary teachers.

In an investigation of the response correlates of teacher stress Kyriacou and Sutcliffe (1979a) found that the correlation between self-reported teacher stress and job satisfaction was negative, while the correlations between stress and intention to leave teaching, and between stress and total days absent were positive. Similarly, Otto (1980) found that job satisfaction and stress scores were inversely related and that teachers' intention to stay or leave teaching was significantly related to their level of job satisfaction. In some studies, teachers have reported illness which they have attributed to the stress of their job (Needle *et al.*, 1981) and this is supported by research which has demonstrated correlations between stress and symptoms (Kyriacou and Sutcliffe, 1978b; Otto, 1980, 1982). Since burnout is regarded as being a reaction to prolonged occupational stress, it is likely that the responses which have been shown to be correlated to self-reported teacher stress will also be related to perceptions of burnout.

The purpose of the present study is to identify biographical characteristics, work pattern, psychological and response variables which distinguish between two groups of teachers experiencing high and low levels of burnout respectively. The biographical variables examined in the study comprise several variables included in previous research, and the work pattern variables investigated are school type, time-fraction worked and position in school. The psychological variables investigated are the personality disposition of hardiness, role stress, social support, self-confidence and teachers' pupil control ideologies. The response variables include measures of career satisfaction, health, career intention, absenteeism and use of regressive coping strategies.

METHOD

Socioeconomic status level of schools

The socio-economic status (SES) of government schools was based on the results of a survey carried out by the Education Department of Victoria, Australia, in July 1984 to identify disadvantaged schools which would receive assistance from the Supplementary Grants Programme (*Education Gazette and Teachers' Aid*, 1985). The SES of non-government schools was estimated by considering the level of government subsidy. Of the 16 participating secondary schools, eight were designated as high SES-government and five were categorised as low SES-government. Three schools were non-Roman Catholic independent schools and were classified as high SES-non-government.

Response rates

A total of 1018 anonymous questionnaires were distributed to the 16 participating schools resulting in the return of 750 usable questionnaires. This represents an overall response rate of 74 per cent which compares favourably with the response rates reported by other researchers using samples of teachers (e.g., Kyriacou and Sutcliffe, 1978b; Schwab and Iwanicki, 1982; Laughlin, 1984a, 1984b; Capel, 1987). The response rates for low SES government, high SES government and high SES non-government schools were 77 per cent, 71 per cent and 76 per cent respectively.

Measures

The Maslach Burnout Inventory (MBI) provides a measure of perceived burnout in terms of the three subscales of emotional exhaustion, depersonalisation and personal accomplishment (Maslach and Jackson, 1981). Each of the 22 items is rated for both frequency and intensity. The frequency rating scale ranges from 1 (a few times a year) to 6 (every day) and the intensity rating scale ranges from 1 (very mild, barely noticeable) to 7 (major, very strong). For both the frequency and intensity scales the respondent scores 0 if the feeling or attitude described is never experienced. Item scores are summed to obtain subscale scores. Higher subscale scores on the MBI indicate higher levels of emotional exhaustion, depersonalisation and personal accomplishment. The MBI has been shown to be reliable and valid when used with workers in the helping professions (Maslach and Jackson, 1981). In the present study, the word *student* replaced the word *recipient* and evidence has been presented to substantiate the reliability and validity of the MBI when adapted in this way for use with teachers (Iwanicki and Schwab, 1981).

In keeping with the techniques used by Maslach and Jackson (1981), a principal components factor analysis with varimax rotation was employed to examine the construct validity of the MBI. These results indicate that, when used with teachers, the same factors identified by Maslach and Jackson (1981) were extracted. This is consistent with the results of other studies which have examined the construct validity of the MBI using samples of teachers in America (Iwanicki and Schwab, 1981; Gold, 1985). Internal consistency estimates of the reliability of each subscale of the MBI for teachers were determined using Cronbach's coefficient alpha. The reliability coefficients for the subscales were 0.89 (frequency) and 0.86 (intensity) for emotional exhaustion, 0.71 (frequency) and 0.72 (intensity) for depersonalisation, and 0.81 (frequency) and 0.81 (intensity) for personal accomplishment. Overall, these reliabilities were practically equivalent to those reported by Maslach and Jackson for their research in the helping professions (1981) and to those reported by Schwab and Iwanicki for their research with teachers (1982).

The Regressive Coping Checklist (RCC) provides a measure of the use of regressive coping actions to deal with stress. The eight items specify behaviours which attempt to deny, minimise, or avoid stressful situations and these were derived from existential personality theory. When Kobasa (1982a) presented three raters with a list of regressive (e.g., "withdrew physically from the situation"), transformational (e.g., "re-evaluated my work priorities"), neutral, or indeterminant (e.g., "thought about it") coping strategies, the eight items which form the RCC were all classified as regressive. Kobasa presents evidence to indicate that these regressive coping actions are characteristic of alienated persons (Kobasa, 1982a). Respondents indicate how frequently they engaged in each of the behaviours to deal with stress at work during the previous year on a five-point scale from 0 (never) to 4 (always). Rather than using the phrase "cope with stress" as did Kobasa, the phrase "deal with stress" was used in the present questionnaire. This was done because the word *cope* implies that the person is successfully handling the stressful situation whereas the word *deal* does not. In this sense, "deal" is a more neutral word than "cope". In the present study, the internal consistency (coefficient alpha) of the regressive coping scale was 0.57, which is lower than the internal consistency of 0.74 reported by Kobasa (1982a).

The Role Questionnaire (RQ) provides a measure of role conflict and role ambiguity and was developed by Rizzo *et al.* (1970). Respondents rate each item on a scale from 1 (definitely not true of my job) to 7 (extremely true of my job). The six items of the role ambiguity subscale are worded positively, so these items are reverse scored. Higher subscale scores on the RQ indicate higher levels of role conflict and ambiguity.

Using the responses of teachers participating in this study, the construct validity of the RQ was examined using a principal components factor analysis with oblique rotation. The results supported the two-factor structure of the RQ found in other studies (Rizzo *et al.*, 1970; Szilagyi *et al.*, 1976; Schuler *et al.*, 1977). In the present study, Cronbach coefficient alpha reliabilities were 0.80 for the role ambiguity subscale and 0.82 for the role conflict subscale. These reliabilities closely corresponded to those of 0.85 for the role conflict subscale and 0.86 for the role ambiguity subscale reported by Pierson (1984) for her research with teachers.

The Work Relationship Index (WRI) gives a measure of social support comprising items from the relationship domain of the Work Environment Scale (WES) developed by Moos (1981) which was designed to measure the social climates of different work settings. The relationship domain is measured by the involvement, peer cohesion, and supervisor support subscales of the WES. The involvement subscale assesses the extent to which employees are concerned about and committed to their jobs. The peer cohesion subscale assesses the degree to which employees are friendly to and supportive of one another, and the supervisor support subscale examines the extent to which senior staff are supportive of employees and encourage employees to be supportive of one another.

In the present study some of the WES items were altered to make them more applicable to a sample of Australian teachers working in post-primary schools. For example, the word *supervisor* was replaced by the phrase *senior staff*, the word *employee* was replaced by the word *teacher*, and the word *organisation* was replaced by the word *school*. The WES item, "employees generally feel free to ask for a raise", was an inappropriate item for teachers since their salaries are determined by award rates, so it was replaced with the item, "teachers generally feel free to ask senior staff for help and advice about work matters". Each subscale comprises nine items and teachers are required to decide whether statements are true or false of their school environment. Responses are scored 0 (false) or 1 (true). For each subscale, four of the items are worded negatively so these items are reverse scored. Higher subscale scores indicate higher levels of involvement, peer cohesion and supervisor support.

Internal consistency estimates of the reliability of each subscale of the WRI were determined using Cronbach's coefficient alpha. The reliability coefficients for the subscales were 0.78 for involvement, 0.69 for peer cohesion and 0.77 for supervisor support reported by Moos (1981). In the present study, the subscales were positively intercorrelated. The correlation between the involvement and peer cohesion subscales was 0.69, the correlation between the involvement and supervisor support subscales was 0.52 and the correlation between the peer cohesion and supervisor support subscales was 0.62. These are higher than the intercorrelations of 0.53, 0.47, and 0.50 for involvement-peer cohesion, involvement-supervisor support and peer cohesion-supervisor support reported by Moos.

When the construct validity of the WRI was examined by means of a principal components factor analysis with oblique rotation, little support for a three-factor solution was found. However, when the involvement, peer cohesion and supervisor support subscales were combined to form a unidimensional scale called the Social Support Scale, the reliability coefficient (Cronbach alpha) was found to be 0.89 which compares favourably with the internal consistency of 0.88 reported by Billings and Moos (1982). Higher scores on the Social Support Scale indicate higher levels of perceived social support in the school environment.

The Pupil Control Ideology (PCI) Form, which was developed by Willower *et al.* (1967), provides a measure of teachers' attitudes toward the teacher-student relationship. The Likert scale items of the PCI Form are scored on a five-step scale ranging from "Strongly Agree" to "Strongly Disagree", with two items being reverse scored. Responses are added to obtain a score which ranges from 20 (extremely humanistic ideology) to 100 (extremely custodial ideology). In the present study, the Cronbach coefficient alpha reliability for the Pupil Control Ideology scale was 0.83.

The *Stress Symptoms Scale* is a list of 25 symptoms which previous researchers identified as being related to high levels of stress (Kyriacou and Sutcliffe, 1978b; Otto, 1982). Teachers indicated how frequently they experienced each symptom during the previous year on a four-point scale from 0 (Never) to 3 (Often). Responses to each item were added to obtain a stress symptom score. The Cronbach reliability coefficient for the Stress Symptoms Scale was found to be 0.90.

Scources of Stress. Teachers were presented with twenty stressors which previous studies reported as being significant sources of stress to teachers (Kyriacou and Sutcliffe, 1978b; Otto, 1980, 1982). For each of these stressors, teachers responded to the following two questions: (a) "As a teacher, how *great a source of stress* are these factors to you?" on a five-point scale from 1 (No stress) to 5 (Extreme stress) and (b) "As a teacher, how *frequently* are these factors a source of stress to you?" on a five-point scale from 1 (Never) to 5 (Always).

The *Hardiness Inventory* is made up of items purporting to measure the personality-based construct of hardiness along the dimensions of commitment, control and challenge. The items in this section are from the shortened composite measure of hardiness developed by Kobasa and Maddi (1982). For 21 of the 36 items, respondents rate each item on a scale from 0 (not at all true) to 3 (completely true). Three items on the challenge scale are worded positively so these are reverse scored. For the 11 items from the Internal-External Locus of Control Scale (Rotter, 1966), respondents choose from two statements the one which best reflects their attitude. Item responses are summed to form subscale scores. Higher scores on the commitment, challenge and control subscales are indicative of lower levels of commitment, challenge and internal control respectively.

The construct validity of the shortened hardiness inventory was examined using the responses of teachers participating in this study by employing a principal components factor analysis with oblique rotation. Three factors with eigenvalues greater than two were extracted which accounted for 25.5 per cent of the variance.

Internal consistency estimates of the reliability of each subscale of hardiness were obtained using Cronbach's coefficient alpha. The reliability coefficients for the subscales were 0.75 for commitment, 0.64 for control and 0.45 for challenge.

Kobasa *et al.* (1982) point to high positive intercorrelations between measures of commitment, control and challenge as indicating the interactive aspects of the three components of hardiness. In the present study, the only high correlation was between commitment and control ($r = 0.48, N = 722, P<0.001$). The correlations between commitment and challenge ($r = -0.12, N = 746, P<0.001$) and between control and challenge ($r = 0.06, N = 722, P<0.05$) were low. These low correlations did not support the view that the three components of hardiness should be regarded as "inextricably intertwined aspects that bear a considerable resemblance to each other'' (Kobasa, 1979, p. 369).

Based on the present analyses, there appears to be no justification for combining the commitment, control and challenge aspects into a composite hardiness measure as suggested by Kobasa and Maddi (1982). Therefore, in this study the three aspects of hardiness were treated separately as suggested by Ganellen and Blaney (1984) who found evidence that the three components of hardiness may be differentially related to social support. Similarly, the three components of hardiness may be differentially related to burnout.

Biographic information collected included sex, marital status, age, employment status, time-fraction worked, teaching experience, years of teaching experience at the same school, position in school, level of education, tertiary institution attended, subjects taught, year levels taught and number and ages of any children which the teacher had. Global measures of occupational and non-occupational stress were operationally defined by responses to the questions, "In general, how stressful do you find being a teacher?" and "In general, how stressful do you find your life away from teaching?" respectively on five-point scales from 1 (not at all stressful) to 5 (extremely stressful).

Career satisfaction was operationally defined by responses to the question, "Overall, how satisfied are you with teaching as a job?" on a five-point scale from 1 (very dissatisfied) to 5 (very satisfied). This has been used as a global measure of career satisfaction in other studies (Kyriacou and Sutcliffe, 1977, 1978b; Otto, 1980, 1982).

Career intention was measured by endorsement of one of five items which ranged from high career intention (remaining as a classroom teacher until retirement) to low career intention (leaving teaching for another occupation). This question was used as a measure of career commitment by McArthur (1981). However, in the present study, career commitment is operationally defined by responses to the question, "If you had the opportunity to choose your career again, how likely is it that you would still choose teaching?", on a five-point scale from 1 (very unlikely) to 5 (very likely).

Career importance was measured by responses to the question, "Considering your life in general, *how important to you* is your career as a teacher?" on a seven-point scale from 1 (not at all important) to 7 (extremely important). Similarly, teachers rated the difficulty of teaching at their present school compared to other schools on a seven-point scale from 1 (not at all difficult) to 7 (extremely difficult). Responses to the question, "Given a completely free choice, how likely is it that you would choose to remain on the staff of your present school next year?", on a five-point scale from 1 (very unlikely) to 5 (very likely), provided a measure of each teacher's desire to remain at their present school. Absenteeism was indexed by teachers' estimates of the total number of days they were absent due to their own ill-health during the previous year, in addition to estimates of the frequency of one day, two-day, three-day and more than three day absences.

Self-confidence was measured by the average rating given on a five-point scale from 1 (not at all like me) to 5 (quite accurately describes me) for the six adjectives: assertive, ambitious, confident, energetic, outgoing, and successful. This measure was obtained from the Health and Daily Living Form (HDL) which seeks personal, social, and health-related information (Billings and Moos, 1982). The reliability coefficient (Cronbach coefficient alpha) was found to be 0.79, which was the same as the alpha coefficient reported by Billings and Moos.

RESULTS AND DISCUSSION

General stress score

An overall stress score was obtained by calculating teachers' average stress ratings for the 20 stressor variables. The stress scale was found to have a Cronbach coefficient alpha of 0.91 with a corrected item-total correlation range from 0.41 to 0.62. The Pearson correlation between the global measure of occupational stress and the stress score was 0.61 ($N = 737$, $P<0.001$).

Stressor factors

A principal components factor analysis with oblique rotation was carried out on the 20 stressor variables and three factors were extracted. The eigenvalues were 7.3, 1.7 and 1.2 with associated variance of 36.4, 8.5 and 6.1 per cent. This accounted for 51 per cent of the variance before rotation. In order of the amount of variance accounted for, the three factors were labelled Professional Needs and Recognition, Student Concerns and Time Demands. Similar factors have been identified in other studies (e.g., Cichon and Koff, 1978; Kyriacou and Sutcliffe, 1978a; Needle *et al.*, 1981; Tellenback *et al.*, 1983; Fisher, 1983; Laughlin, 1984a). The factors were positively intercorrelated. The correlation between Factors 1 and 2 was 0.52, the correlation between Factors 1 and 3 was 0.52 and the correlation between Factors 2 and 3 was 0.42.

For each teacher, stressor factor scores were generated for each of the three factors. Using the "complete estimation" method, factor scores were obtained from the summation of terms derived from the standardised variable scores and the factor loadings for each variable in the factor score coefficient matrix.

High and low burnout groups

The median for each of the six subscales of the MBI was computed and these values are shown in Table 1. The high burnout group comprised teachers who were above the median on the emotional exhaustion and depersonalisation subscales and below the median on the personal accomplishment subscales. The low burnout group comprised teachers who were below the median on the emotional exhaustion and depersonalisation subscales and above the median on the personal accomplishment subscales. Adopting these criteria in the formation of high and low burnout groups, there were 81 teachers in the high burnout group (10.8 per cent of all teachers who participated in the study) and 94 teachers in the low burnout group (12.5 per cent).

TABLE 1

MEDIANS FOR THE MBI SUBSCALES

Variable	Frequency	Intensity
Emotional Exhaustion	21.97	31.97
Depersonalisation	6.13	8.98
Personal Accomplishment	33.98	37.95

Maslach and Jackson (1981) assert that the higher the degree of experienced burnout, the higher the scores on the emotional exhaustion and depersonalisation subscales and the lower the scores on the personal accomplishment subscale. They defined a score as reflecting a high level of burnout if it was in the top third of their normal distribution for emotional exhaustion and depersonalisation and in the bottom third for personal accomplishment. Using the scores from in the MBI Manual (1981) as the criteria for measuring a high level of burnout, 3.7 per cent of the teachers in the present study experienced a high level of burnout and this proportion is similar to the results reported by Beasley *et al.* (1983).

Biographical and work pattern variables

Using the high and low burnout groups formed by median splits as described above, a series of chi-square tests was carried out to examine whether the composition of the high and low burnout groups differed with respect to the variables of school type, sex, marital status, having children, age, teaching experience, teaching experience at present school, position in school, subjects taught, year levels taught, teaching year 12 and level of education.

The results indicated that the composition of the high and low burnout groups did not differ in terms of sex (χ^2 (1) = 0.012, P>0.05), marital status (χ^2 (1) = 1.13, P>0.05), having children (χ^2 (1) = 0.443, P>0.05), age (χ^2 (3) = 5.92, P>0.05), subjects taught (χ^2 (2) = 0.582, P>0.05), year levels taught (χ^2 (6) = 9.554, P>0.05), teaching year 12 (χ^2 (1) = 0.012, P>0.05) or level of education (χ^2 (1) = 2.875, P>0.05). These results are in accordance with Otto's (1982) finding that the proportion of highly stressed teachers remained constant regardless of age level or subjects taught.

In the present research, there were differences between the high and low burnout groups according to school type (χ^2 (2) = 112.154, P<0.01), teaching experience (χ^2 (3) = 98.271, P<0.05), teaching experience at present school (χ^2 (3) = 11.18, P<0.05) and position in school (χ^2 (2) = 10.249, P<0.01). The proportion of teachers in the high burnout group from low SES government schools (45.7 per cent) was higher than in the low burnout group (26.6 per cent). In contrast, the proportion of teachers in the high burnout group from high SES non-government schools (8.6 per cent) was less than in the low burnout group (26.6 per cent). These results are consistent with Pratt's (1978) finding that primary teachers working in low SES areas experienced higher levels of stress than teachers working in higher SES areas. Teachers in the low burnout group tended to have more overall teaching experience, and had taught at their present school for longer periods of time than teachers in the high burnout group. Similar results were reported by Beasley *et al.* (1983) who found that teachers who scored highly on the MBI had six years or less of teaching experience.

In the present study, more teachers involved mainly in classroom teaching duties were in the high burnout group (65.4 per cent) compared to the low burnout group (41.5 per cent), whereas more teachers with additional responsibilities were in the low burnout group (46.8 per cent) compared to the high burnout group (25.9 per cent). These results contradict Otto's (1982) finding that about the same proportion of highly stressed teachers were among those with or without co-ordinating responsibilities and among those with many or few years of teaching experience. However, the results of the two studies are not directly comparable. The aim of the present analysis was to compare the responses of teachers who perceived themselves as experiencing high and low levels of burnout, whereas Otto's (1982) aim was to compile a profile of the most highly stressed teachers in her study. Apart from the different objectives of the two studies, it should be noted that stress and burnout are not synonymous.

Sources of stress

A series of t-tests was conducted to compare the scores of teachers in the high and low burnout groups on the 25 variables listed in Table 2.

From Table 2, it can be seen that there was no significant difference between teachers in the high and low burnout groups on the global measure of non-occupational stress. This is consistent with Otto's (1982) assertion that, for the majority of teachers, susceptibility to

TABLE 2

COMPARISONS BETWEEN HIGH AND LOW BURNOUT GROUPS

Variable	High Burnout		Low Burnout		df	t
	Mean	N	Mean	N		
Global occupational stress	3.70	81	2.15	94	173	14.16***
Global non-occupational stress	2.05	81	1.92	93	172	1.03
Stress attributed to teaching	3.00	80	1.85	93	171	8.94***
Stress Score	1 79	80	0.75	94	172	11.52***
Prof. Needs and Recognition	1.46	74	0.58	88	160	8.60***
Student Concerns	2.22	74	0.95	88	160	12.24***
Time Demands	2.07	74	1.28	88	160	7.70***
Commitment	8 07	81	2.68	93	172	9.33***
Control	11.36	80	7.99	92	170	6.36***
Challenge	14.84	81	16.01	93	172	-2.66**
Social Support	14.19	78	21.09	94	170	-7.84***
Role Conflict	34.11	81	24.91	94	173	6.01***
Role Ambiguity	23.20	81	14.40	94	173	9.80***
Pupil Control Ideology	60.28	81	52.64	94	173	4.75***
Satisfaction	2.65	81	4.31	94	173	-11.06***
Career Intention	3.00	80	2.26	90	168	3.86***
Career Likelihood	2 49	81	3.82	92	171	-7.23***
Career Importance	4.21	81	5.73	94	173	-8.05***
Self-Confidence	3.04	80	3.96	94	172	-9.26***
Physical Health	33 63	81	13.77	94	173	11.60***
Short Term Absences	4.30	79	2.47	94	171	4.64***
Total Days Absent	11.72	78	4.78	92	168	3.55***
Regressive Coping	11.49	80	5.04	94	172	10.45***
Staying at Present School	3.00	81	4.29	94	173	-5.91***
Difficulty of School	3.92	79	2.32	92	169	6.23***

** $P < 0.01$, *** $P < 0.001$

work stress is not increased by stress in their private lives. In contrast, the results are inconsistent with Laughlin's (1984a) finding that psychological well-being, as measured by the General Health Questionnaire, was strongly influenced by undesirable life events. Except for the global measure of non-occupational stress, there were significant differences between the high and low burnout groups for all the other variables in Table 2. In comparison to teachers in the low burnout group, teachers in the high burnout group recorded higher scores on the global measure of occupational stress, attributed more of the total stress in their lives to teaching, recorded higher general stress scores and perceived more stress from the stressor factors, Professional Needs and Recognition, Student Concerns and Time Demands.

Teachers in the high burnout group rated Student Concerns as being the factor which caused them the most stress whereas teachers in the low burnout group rated Time Demands as being the factor which concerned them the most. Beasley *et al.* (1983) found that high MBI responders ranked Parent/Teacher Relations as being the most stressful factor, which was different from the total sample which ranked Time Management problems as being the most stressful factor. These results suggest that teachers who experience high levels of burnout find stressors associated with interpersonal relationships to be more stressful than teachers who experience low levels of burnout. This is not surprising considering that one of the characteristics of burnout is a depersonalised attitude toward others.

Psychological and response variables

Overall, teachers in the high burnout group were less hardy than those in the low burnout group as evidenced by lower levels of commitment and a more external locus of control. The finding that hardy teachers experienced lower levels of burnout supports the contention that hardiness acts as a stress resistance resource to moderate the negative effects of stress (e.g., Kobasa, 1979, 1982a). Teachers in the high burnout group perceived less social support in their school environment than teachers in the low burnout group. This finding is in agreement with the literature which contends that social support helps to reduce the negative effects of stress (e.g., Pinneau, 1976; House, 1981). In particular, these results support Laughlin's (1984a) finding that teachers' perceptions of stress were reduced by perceived support from colleagues and special friends on their school staff and that poor psychological health was attenuated by colleague support. These results suggest that teachers who work in an environment which they perceive as being supportive are less likely to experience high levels of burnout.

Teachers in the high burnout group perceived more role stress than teachers in the low burnout group as indicated by higher levels of role conflict and role ambiguity. This is consistent with the results reported by other researchers who have found a positive association between role stress and burnout (Westerhouse, 1980; Crane, 1982; Schwab and Iwanicki, 1982; Mansfield, 1984).

Teachers in the high burnout group had more custodial PCIs than those in the low burnout group. This is consistent with the results of a recent study which found that teachers with custodial PCIs experienced more teaching anxiety than teachers with less custodial PCIs (Docking, 1985).

Compared to teachers in the low burnout group, teachers experiencing high levels of burnout were less satisfied with their career, were absent more frequently, had significantly poorer physical health and reported less career commitment as indicated by lower levels of career intention and career likelihood. These findings are in agreement with the result of other studies which have examined some of these response correlates of teacher stress (Kyriacou and Sutcliffe, 1979a; Otto, 1980).

In addition to confirming some of the results of previous research, the analyses highlight other differences between teachers who experience high and low levels of burnout. For instance, teachers in the high burnout group perceived themselves as less self-confident and rated their career as being less important to them than teachers in the low burnout group. This supports the view that there is a negative relationship between burnout and self-concept (Schlachter, 1981; Blough, 1981). Teachers in the high burnout group used regressive coping strategies more frequently than teachers in the low burnout group and this is consistent with the results reported by Kobasa (1982a). For example, when studying stress resistance among lawyers, Kobasa found that increases in strain symptoms were related to increases in frequency of use of regressive coping strategies.

Overall, teachers in the high burnout group rated their school as being more difficult to teach at and were less likely to want to remain on the staff of their present school if they were given a free choice when compared to teachers in the low burnout group. These findings indicate other response variables which are related to perceptions of burnout.

Models of burnout

In order to investigate the relative importance of the background, work pattern and psychological variables in accounting for the variance in each of the three burnout subscales, stepwise multiple regression analyses were used to obtain a model for each subscale. Table 3 shows that lower levels of commitment and control, higher levels of role conflict and role ambiguity and teaching at lower SES schools contributed to more frequent feelings of emotional exhaustion. More frequent feelings of depersonalisation were related to low levels of commitment, high levels of role conflict and role ambiguity, a more custodial pupil control ideology and younger teachers.

The analysis showed that higher levels of role ambiguity contributed to less frequent feelings of personal accomplishment, while higher levels of commitment and role conflict

TABLE 3

STEPWISE MULTIPLE REGRESSION ANALYSIS FOR THE RELATIONSHIPS OF BIOGRAPHICAL, WORK PATTERN AND PSYCHOLOGICAL VARIABLES TO BURNOUT (N=667)

Step	Variable(s) Entered	Beta at Entry	Final Beta	Increment in R^2	F to Enter
		Emotional Exhaustion (Frequency)			
1.	Commitment	0.4480	0.2594	0.2007	167.01***
2.	Role Conflict	0.2913	0.2086	0.0784	72.25***
3.	Role Ambiguity	0.1664	0.1493	0.0219	20.77***
4.	Control	0 1309	0.1346	0.0128	12.38***
5.	School Type		-0.0901	0.0074	7.25**
				$R^2 = 0.3212$	$\hat{R}^2 = 0.3161$
		Depersonalisation (Frequency)			
1.	Commitment	0.4229	0.2946	0.1789	144.86***
2.	Role Conflict	0.1953	0.1817	0.0353	29.79***
3.	PCI	0.1802	0.1793	0.9310	27.22***
4.	Age	-0.1051	-0.0971	0.0110	9.76**
5.	Role Ambiguity		0.1003	0.0079	7.09**
				$R^2 = 0.2641$	$\hat{R}^2 = 0.2585$
		Personal Accomplishment (Frequency)			
1.	Role Ambiguity	-0 3406	-0.3081	0.1160	87.24***
2.	Commitment	-0.2080	-0.2308	0.0376	29.47***
3.	Role Conflict		0.1397	0.0164	13.12***
				$R^2 = 0.1700$	$\hat{R}^2 = 0.1662$
		Emotional Exhaustion (Intensity)			
1.	Role Conflict	0.4177	0.2789	0.1745	140.54***
2.	Commitment	0.2759	0.1701	0.0704	61.88***
3.	Role Ambiguity	0.1658	0.1576	0.0217	19.64***
4.	Control		0.1365	0.0139	12.83***
				$R^2 = 0.2805$	$\hat{R}^2 = 0.2762$
		Depersonalisation (Intensity)			
1.	Commitment	0.3796	0.2899	0.1441	111.95***
2.	Role Conflict	0.2027	0.1984	0.0380	30.83***
3.	PCI	0.1327	0.1253	0.0168	13.91***
4.	Children	0.0981	0.1045	0.0096	8.02**
5.	Sex		-0.0941	0.0086	7.24**
				$R^2 = 0.2171$	$\hat{R}^2 = 0.2112$
		Personal Accomplishment (Intensity)			
1.	Role Ambiguity	-0 2678	-0.2835	0.0717	51.36***
2.	PCI	-0.1992	-0.1561	0.0397	29.63***
3.	Role Conflict	0.1995	0.2306	0.0344	26.67***
4.	Commitment		-0.1792	0.0260	20.75***
				$R^2 = 0.1718$	$\hat{R}^2 = 0.2762$

** $P < 0.01$, *** $P < 0.001$

contributed to more frequent feelings of personal accomplishment. Higher levels of role conflict and role ambiguity and lower levels of commitment and control contributed to more intense feelings of emotional exhaustion. Lower levels of commitment, higher levels of role conflict and a more custodial pupil control ideology contributed to more intense feelings of depersonalisation. Teachers with children and female teachers experienced less intense feelings of depersonalisation than their childless and male counterparts respectively. Higher levels of role ambiguity and lower levels of commitment were related to less intense feelings of personal accomplishment, while a less custodial pupil control ideology and higher levels of role conflict contributed to more intense feelings of personal accomplishment.

CONCLUSIONS

The study has identified a number of variables which are related to feelings of burnout. The results indicated that hardiness and role stress are the most significant predictors of the three aspects of burnout. Further, the analyses showed that teacher biographical characteristics (sex, age, marital status, teaching experience, level of education and parental status) are less significant predictors of burnout than work pattern variables (school type, position in school, time-fraction worked) and psychological variables (hardiness, social support, role stress, pupil control ideology). However, the findings do not enable conclusions to be made about causality. For example, the multiple regression analyses indicated that commitment accounted for a significant amount of variance on all of the burnout subscales and that teachers in the high burnout group were less committed than teachers in the low burnout group. From these results, it can be concluded that there is a relationship between levels of commitment and feelings of burnout. But it cannot be concluded that a low level of commitment causes high levels of burnout. Instead, it may be the experience of high levels of burnout which leads some teachers to develop reduced levels of commitment. Longitudinal research would be able to examine these relationships by monitoring changes in variables such as commitment, pupil control ideology and feelings of burnout. In addition, such research would facilitate an understanding of the complex processes between feelings of burnout and the other variables examined in this study.

The results of this study showed that school type was related to perceptions of stress and burnout. The finding that teachers working in low SES government, high SES government and high SES non-government schools reported different levels of burnout indicates that future research might examine the reasons for these differences. Research comprising in-depth case studies of schools would help to pin-point the reasons for these differences and identify the parameters related to the responses of teachers working in different types of schools.

The present research indicates that teachers who experience high levels of burnout use regressive coping strategies more frequently than teachers who experience low levels of burnout. Regressive coping strategies are basically means of escaping from stressful situations or denying that the situation is stressful. They are not strategies which attempt to appraise and resolve the stressful aspects of the situation which a person confronts. On the basis of the present research, it cannot be said whether teachers are more likely to experience high levels of burnout because they use regressive coping strategies or whether they are more likely to use regressive coping strategies because they have experienced high levels of burnout. Since regressive coping strategies are generally regarded as being ineffective ways to deal with stressful situations (Kobasa, 1982a), it would be useful to know more about the circumstances which contribute to their development, so that appropriate intervention programmes might be initiated. Studies using longitudinal designs and qualitative data are needed to complement this domain of educational research.

Correspondence and requests for reprints should be addressed to Dr G. N. Molloy, Faculty of Education, Monash University, Clayton, Melbourne, Victoria 3168, Australia.

REFERENCES

Beasley, C. R., Myette, B. M., and Serna, B. (1983). On-the-job stress and burnout: contributing factors and environmental alternatives in educational settings. Paper presented at the Annual Meeting of the American Educational Research Association (Montreal, Canada, April 11-15).

Billings, A. B., and Moos, R. H. (1982). Work stress and the stress-buffering roles of work and family resources. *J. occup. Behav*, 3, 215-232.

Blough, S. G. (1981). Perceived stress as a function of setting and personality. *Diss. Abstr. Int.*, 41, 4285A.

Capel, S. A. (1987). The incidence of and influences on stress and burnout in secondary school teachers. *Br. J. educ. Psychol.*, 57, 279-288.

Carlson, R. P. (1964). Environmental constraints and organizational consequences: the public school and its clients. In Griffiths, D. E. (Ed.), *Behavioural Science and Educational Administration*. 63rd Yearbook of the National Society for the Study of Education. Chicago: University of Chicago Press.

Cichon, D., and Koff, R. J. (1978). The teaching events stress inventory. Paper presented at the annual meeting of the American Educational Research Association, Toronto, Canada.

Coates, T. J., and Thoresen, C. E. (1976). Teacher anxiety: a review with recommendations. *Rev. educ. Res.*, 46 (2), 159-184.

Cooper, C. L., and Marshall, J. (Eds.) (1980). *White Collar and Professional Stress*. Chichester: John Wiley and Sons.

Crane, S. (1982). An analysis of the relationship among personal and professional variables, role ambiguity, role conflict, and perceived burnout of special education teachers. *Diss Abstr. Int.*, 42, 3549A.

Docking, R. A. (1985). Changing teacher pupil control ideology and teacher anxiety. *J. Educ Teaching*, 11 (1), 63-76.

Docking, R. A., and Docking, E. (1984). Reducing teacher stress. *Unicorn*, 10 (3), 261-274.

Dunham, J. (1980). An exploratory study of staff stress in English and German comprehensive schools. *Educ. Rev.*, 32 (1), 11-20.

Education Gazette and Teachers' Aid (1985). 13 June, pp. 393-396.

Fisher, M. (1983). *Stress and Burnout in Northern Territory Teachers: An Exploratory Survey*. Northern Territory Department of Education.

Fisher, M. (1984). Teacher burnout: some Australian findings. *Unicorn*, 10 (3), 226-232.

Ganellen, R. J., and Blaney, P. H. (1984). Hardiness and social support as moderators of the effects of life stress, *J. Person. soc. Psychol.* 47 (1), 156-163.

Gold, Y. (1985). The relationship of six personal and life history variables to standing on three dimensions of the Maslach Burnout Inventory in a sample of elementary and junior high school teachers. *Educ. psychol. Measur.*, 45, 377-387.

House, J. S. (1981). *Work Stress and Social Support*. Reading, Mass.: Addison-Wesley.

Iwanicki, E. F., and Schwab, R. L. (1981). A cross-validation study of the Maslach Burnout Inventory. *Educ. psychol Measur.*, 41 (4), 1167-1174.

Kobasa, S. C. (1979). Stressful life events, personality, and health: an inquiry into hardiness. *J. Person. soc. Psychol.*, 37, 1-11.

Kobasa, S. C. (1982a) Commitment and coping in stress resistance among lawyers. *J. Person soc. Psychol.*, 42 (4), 707-717.

Kobasa, S. C. (1982b). The hardy personality: toward a social psychology of stress and health. In Sanders, G. S. and Suls, J. (Eds.), *Social Psychology of Health and Illness*. Hillsdale, N.J. Lawrence Erlbaum Associates.

Kobasa, S. C., and Maddi, S. R. (1982). *Hardiness Measurement*. Memo received from Department of Behavioural Sciences, University of Chicago.

Kobasa, S. C., Maddi, S. R., and Kahn, S. (1982). Hardiness and health: a prospective study. *J. Person soc. Psychol.* 42 (1), 168-177.

Kyriacou, C., and Sutcliffe, J. (1977). Teacher stress: a review. *Educ. Rev.*, 29 (4), 299-306.

Kyriacou, C. and Sutcliffe, J. (1978a). A model of teacher stress. *Educ. Stud.*, 4 (1), 1-6.

Kyriacou, C. and Sutcliffe, J. (1978b). Teacher stress: prevalence, sources, and symptoms. *Br J. educ. Psychol.*, 48, 159-167.

Kyriacou, C, and Sutcliffe, J. (1979a). Teacher stress and satisfaction. *Educ. Res.*, 21 (2), 89-96.

Kyriacou, C, and Sutcliffe, J. (1979b) A note on teacher stress and locus of control *J. occup. Psychol.*, 52, 227-228.

Laughlin, A. (1984a). Occupational stress and its relationship to social supports and life turbulence. Unpublished doctoral dissertation, University of New South Wales, Australia.

Laughlin, A. (1984b) Teacher stress in an Australian setting: the role of biographical mediators. *Educ. Stud.*, 10 (1), 7-22.

Mansfield, J. R. (1984). Role conflict role ambiguity, powerlessness and burnout among high school department heads. *Diss Abstr Int.*, 44, 1996A

Maslach, C (1976). Burned-out. *Hum. Behav.*, 5 (9), 16-22.

Maslach, C, and Jackson, S. E. (1981). *The Maslach Burnout Inventory*. Palo Alto, CA: Consulting Psychologists Press.

McArthur, J. (1981). *The First Five Years of Teaching*. ERDC Report No. 30, Australian Government Publishing Service, Canberra.

Moos, R. H. (1981). *Work Environment Scale Manual*. Palo Alto, CA. Consulting Psychologists Press.

Needle, R. H., Griffin, T., and Svendsen, R. (1981). Occupational stress: coping and health problems of teachers. *J School Health*, 5 (3), 175-181.

Otto, R. (1980). Structural sources of teacher stress in state high schools. La Trobe University Sociology Papers.

Otto, R. (1982). Occupational stress among teachers in post-primary education: a study of teachers in technical school and some comparative data on high school teachers. Department of Sociology, La Trobe University.

Otto, R. (1986). *Teachers Under Stress*. Hill of Content, Melbourne.

Pierson, D. A. (1984). Perceived intensity of burnout among supportive service groups and a comparison teacher group as related to perceived need deficiencies, perceived role stress, and selected personal and professional variables and percepts. *Diss. Abstr. Int.*, 44, 3018A.

Pinneau, R. S. (1976). Effects of social support on psychological and physiological stress. *Diss. Abstr. Int.*, 36, 53598-53560B.

Pratt, J. (1978). Perceived stress among teachers: the effects of age and background of children taught. *Educ. Rev*, 30 (1), 3-14.

Rizzo, J., House, R., and Lirtzman, S. (1970). Role conflict and ambiguity in complex organizations. *Admin. Sci. Q*, 15, 150-163.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychol Monogr.*, 80 (1), 1-28.

Schlachter, P. (1981). Concepts of self and perceived stress: a study of correctional officers. *Diss. Abstr. Int*, 41, 4294A.

Schuler, R. S., Aldag, R., and Brief, A. (1977). Role conflict and ambiguity: a scale analysis. *Organiz. Behav Hum. Perform*, 20, 119-128.

Schwab, R L., and Iwanicki, E. F. (1982). Perceived role conflict, role ambiguity, and teacher burnout. *Educ. Admin Q.*, 18 (1), 60-74.

Selye, H. (1976). *The Stress of Life* (Rev. edit.) New York: McGraw-Hill.

Szilagyi, A., Sims, H., and Keller, R. (1976). Role dynamics, locus of control, and employee attitudes and behaviour. *Academy of Management J.*, 19, 259-275.

Tellenback, S., Brenner, S., and Lofgren, H. (1983). Teacher stress: exploratory model building. *J. occup. Psychol*, 56, 19-33.

Westerhouse, M. A. (1980). The effects of tenure, role conflict and role conflict resolution on the work orientation and burn-out of teachers. *Diss Abstr. Int*, 41, 174A-175A.

Willower, D. J, Eidell, T. L., and Hoy, W. K. (1967). *The School and Pupil Control Ideology*. Penn State Studies Monograph 24, Pennsylvania State University.

(*Manuscript received* 3rd *January*, 1989)

UNIVERSITY COLLEGE LONDON
Department of Psychology

DIPLOMA IN FAMILY THERAPY

Applications are invited for this course aimed at training practitioners to work with families in a multi-cultural society. The one-year course consists of one and a half days per week formal tuition, which includes theory and reading seminars, live supervision of family work, videotape supervision and skills training.

The Course Director is Dr Alan Cooklin, and the teaching staff will be drawn from the Marlborough Family Service, from University College London, and some visiting staff.

Applicants must have a relevant professional qualification and must have completed some recognised introductory work in Family Therapy.

Further information and application forms from:
Dr Rosine Perelberg, Course Organiser, Marlborough Family Service,
38 Marlborough Place, LONDON NW8 0PJ
Telephone 01-624 8605

Closing date: 31st May 1990

Br. J. educ. Psychol., 60, 52-62, 1990

VARIETIES OF CURRICULAR EXPERIENCE: BACKMARKERS AND FRONTRUNNERS IN THE PRIMARY CLASSROOM

BY JOHN INGRAM AND NORMAN WORRALL
(*University of London Institute of Education*)

SUMMARY. In the first study a class of 10-year-olds was observed for the first and last five weeks of the autumn term. On both a class and individual child basis there was considerable difference between teacher-believed and child-actual timetables, and this increased over the term. In the second study, 53 children were classified into "frontrunners" or "backmarkers", depending on success in pacing the taught curriculum. For backmarkers, strategies for maintaining contact were mainly child-determined, though for frontrunners mainly teacher-determined. Although the teachers, who proved to be traditionally directive, believed their paper timetable closely represented children's true classroom experience, a picture emerged of most children correctly perceiving themselves as backmarkers, and as having responsibility for instigating their own, largely passive, coping strategies with subsequent negative feelings about themselves and their situation.

INTRODUCTION

IN the past ten years there has been a shift from dependence on information from teachers for a construction of classroom reality to a more balanced perspective where both pupil-teacher and pupil-pupil processes can be incorporated. A notable influence in North America has been Doyle's argument for seeing the classroom not as a vacuum in which cognitive processes occur, but as a public setting for a complex, social group within which events are happening simultaneously, often unpredictably and with little time for reflection (Doyle, 1986).

In the United Kingdom, the "new look" serves to catch the ethnographic mood of current educational research. Ironically, it also at last makes contact with the recommendations of the Plowden Report, now 20 years old, which urged that children should be more active agents in the learning process. The work of Deci and Ryan (1986), Ingram and Worrall (1987), Bennett and his associates (1984), Desforges and Cockburn (1987) and of Galton (1987b) has broken with exclusive dependency on teachers' viewpoints and looked at differences between teacher intentions and pupil experiences.

The line of research taken by Bennett and by Desforges has focused on contrasts within a particular curriculum subject or even within a task, as in Desforges' and Cockburn's (1987) finding of marked between-child variability in exposure to a supposed fixed set of mathematical concepts. Taking a broader, across-curriculum perspective, a recent study in 33 London primary schools by Farquhar and associates (1987) found that not only did teachers differ widely in their views and their practices concerning an appropriate curriculum for top infants, but that there were marked discrepancies between the teacher-intended curriculum, as being taught to the class as a whole, and the curriculum as received by individual children.

The present paper reports two studies which are also concerned with discrepancies across the curriculum as a whole, but which instead of depending on teacher evidence focus on the contrast between teacher reports and reports from the children themselves, as validated by an observer.

The first study, which served to provide baseline data, followed a class of 12 children over 10 weeks. The children had a formal-directive teacher and the findings are complementary to those of Bennett *et al.* (1984) and Desforges and Cockburn (1987) in illustrating the kinds of mismatch that can build up between what the teacher believes is being taught, and what children themselves experience.

The second study examined strategies used by children out of step with the lesson being taught, and provides evidence of feelings as a result of being behind or ahead of the rest of the class. The study involved interviews with a larger number of children, and attempted to capture the texture of the child's experience, as achieved in the later Oracle studies (Galton, 1987b).

STUDY ONE: TEACHER-CHILD DISCREPANCIES IN PERCEPTION OF THE TAUGHT CURRICULUM

METHOD

The class

The children were third-year juniors, six boys and six girls aged 9 to 10 years. The class was selected because their teacher was a good example of the formal, directive teaching style — in fact, very close to Bennett's (1976) "Style 12" (non-integrated approach to the classroom; separate subject lessons; prescribed classroom seating; central blackboard; general curb on child movement or talking; highly directive in all activities). This characterisation is based on two days' observation by one author (JI), together with the teacher's own description of his style and the corroboration of colleagues who were familiar with his classroom approach.

Procedure

Data were collected daily over two five-week periods, one period at the beginning and one at the end of the first term of the school year.

At the outset, the teacher made available a timetable showing the planned curriculum for the two data collection periods.

A grid record was then kept by each pupil listing curriculum activities during each morning and afternoon. Pupils were told simply to write a note in the box indicating what they had been doing. In practice even the laconic notes often given when an activity was continued were capable of interpretation. (For example, set against the previous entry, "Mosque" could only mean that the boy had continued to write his story about going to the mosque.) Grid entries were completed immediately before going to lunch and before leaving for home in the afternoon.

A similar grid record was kept by the teacher, again itemising the curriculum areas that had been covered that morning and afternoon.

There was also a spot-check by an observer of the curricular activities in which the child was, in fact, involved during the morning and afternoon. Thus the observer's grid could be placed against child and teacher grids as a validity estimate for the data as a whole. These validity spot-checks were carried out across all 12 children every day, morning or afternoon, and took about half-an-hour. While not as comprehensive as a continuous observational record (Farquhar *et al.*, 1987), no error artefact was expected, since preliminary observation had shown that children of this age, as well as the teacher, were making categorisations that almost always coincided with those of the observer.

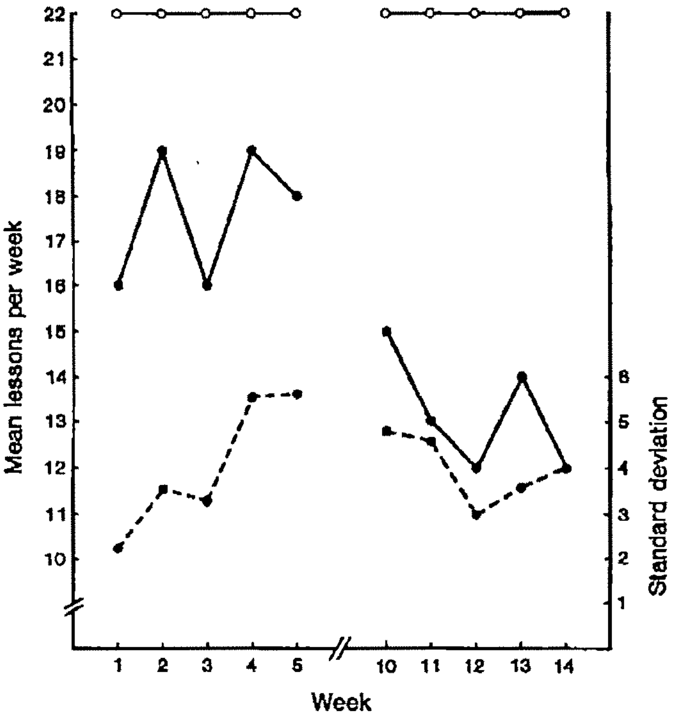
Each evening the various data were collated and comparative profiles progressively assembled for the two five-week periods.

RESULTS

Figure 1 gives overall data for the two five-week blocks, and provides some general context for more detailed findings. The broken-line graph shows the teacher's record of lessons as planned and as actually taught. This was a constant 22 lessons per week (excluding PE, swimming and games), with an identical fit between the "planned" and the "as taught" timetable, except for slight disruption from Christmas activity rehearsals on four out of the 50 measurement days. The solid line graphs are the children's records. They show how the number of lessons recorded by the class as a whole varies both over time (weekly means) and among children within a given week (standard deviations). It can be seen that between the first and last five-week blocks (a) there is a widening *discrepancy* between the teacher- and child-recorded mean number of lessons for the curriculum as a whole; (b) there is an overall *decrease* in the number of curricular units as recorded by children, from 17.21 to 13.14 per week; (c) there is an increase in the *variability* of individual child experiences over the first five-week block, with a return to more "average" levels in the last weeks of term.

FIGURE 1

Number of Lessons Recorded by the Class as a Whole for the First and Last Five Weeks of the Autumn Term.



Note: Solid-line graphs are means from children's records, and dotted graphs the standard deviations. The open-circle graph shows the teacher's record — a constant 22 lessons per week.

Thus at this general level of the class and the curriculum as a whole, there seems to be only a loose relationship between the teacher- and pupil-recorded curriculum, with this looseness being far from self-correcting over the first term. It should also be noted here that the validity spot-checks referred to in the Method, and carried out on a time-sampling basis, gave full credence to the children's records as being accurate.

TABLE 1

BREAKDOWN OF EACH CHILD'S CURRICULAR EXPERIENCE FOR THE CASE OF MATHEMATICS AND ENGLISH

(a)	Total Maths Periods as Taught	Maths Periods Experienced by Each Child (1 - 12)												
		1	2	3	4	5	6	7	8	9	10	11	12	Totals
First 5 Weeks	25	28	40	36	39	34	26	24	26	26	22	18	19	338
Last 5 Weeks	25	38	23	26	17	19	23	35	22	21	23	19	16	282
Change Over Term	0	+10	-17	-10	-22	-15	-03	+11	-04	-05	+01	+01	-03	

(b)	Total English Periods as Taught	English Periods Experienced by Each Child (1 - 12)												
		1	2	3	4	5	6	7	8	9	10	11	12	Totals
First 5 Weeks	20	14	18	15	04	12	10	06	10	09	04	11	05	118
Last 5 Weeks	20	03	17	02	00	15	02	01	02	04	03	01	09	59
Change Over Term	0	-11	-01	-13	-04	+03	-08	-07	-08	-05	-01	-10	+04	

Individual differences in curriculum as experienced

Table 1 (a) and 1 (b) provide greater resolution by taking the major cases of mathematics and English, and breaking the data down on an individual child basis.

Mathematics was the curriculum subject which actually had the best fit between teacher and child records. Nevertheless, it can be seen in Table 1 (a) that over all 10 weeks, as well as within each five-week block, there is considerable between-pupil variation in mathematics exposure, standard deviation = 7 periods, as well as considerable deviation from the figure assumed by the teacher of 25 units of teaching. In the first five-week block the range of recorded exposure is from 18 to 40 periods (or in terms of per-week means, from 3.6 to 8.0) and in the second block there is a similar range of from 16 to 38 periods. Moreover, there is no consistency for individual children between the two five-week blocks, and for some a considerable discrepancy: for example, children 2, 4 and 5 show large decreases, while children 1 and 7 show large increases.

Table 1 (b) shows parallel data for the second major curricular subject, English. The teacher-assumed number of lessons over the first five-week block was 20 (4 per week), while the mean child-recorded figure was 10.60, range 4 to 18. Thus, no child recorded the intended quota: in the last five-week block in particular, the table shows that only child 2 and child 5 recorded anything approaching 20 English lessons.

DISCUSSION

All teachers obviously expect variation and hence some discrepancy between what they intend and what they can in fact achieve, as made clear by the Farquhar *et al.* (1987) study. However, this has not been the present issue. A single class has been followed for 10 weeks and has yielded data showing a non-correcting, poor fit between the teacher's own *believed* curriculum and the curriculum as actually recorded by the children in that class.

This discrepancy was evidenced in several ways: (a) except for the case of mathematics, the children were recording fewer lessons than the teacher supposed; (b) rather than being susceptible to natural correction, the discrepancy had increased by the last five-week period; (c) for the elected examples of mathematics-and English, there were large between-pupil variations in records; (d) patterns for given pupils over the two five-week periods, showed wide inconsistencies.

The particular pattern in this first study might of course be no more than indicative of classroom practices, in the sense that the teacher might have been ambiguous in signalling change of lesson. This seems unlikely, since when there was no natural end to a lesson, as at assembly, playground break or lunch, the teacher's consistent practice was to say, "Alright, stop now and put your books away", and monitor class response before changing the lesson.

Nor are the present results likely to be explainable in terms of peculiar characteristics of this one teacher, since the findings of Farquhar *et al.* (1987, Table 9) are at least suggestive that a similar kind of discrepancy is more widespread. In their study, when teachers recognised that a discrepancy was occurring, about half cited as reasons dispositional factors such as motivation, co-operation or concentration, and half again cited children's entry skills or ability. Attendance patterns and home influences were cited by considerably fewer teachers. None of the teachers seemed to think the reason might be their own classroom practices. Such implicit corroboration of the kinds of findings emerging from this first study encouraged us to look to a larger group of children and their teachers to discover more about what processes might be operating in the definition of classroom work.

STUDY TWO: CHILDREN'S STRATEGIES FOR KEEPING IN STEP WITH THE TAUGHT CURRICULUM

The aim of the second study was to get down to the level of the individual child's experience of the curriculum, and obtain interview data from children that might be set alongside the observational data of, for example, the later Oracle studies (Galton, 1987b). Galton's observational data defined interesting pupil types, such as the intermittent versus the solitary workers, with the latter group further appearing to consist of children who either worked conscientiously at set tasks (the "hard grinders") or who worked with studied deliberation at subterfuge tasks such as breaking and resharpening pencil points or measuring and remeasuring lines (the "easy riders").

It was decided to take a new and larger sample of children, and explore both opinions and feelings on the kind of discrepancies emerging in the first study, and in particular to discover what strategies children used to maintain contact with the curriculum as it was being taught.

Pupil interviews

(i) The interview questions to be used were developed from open-ended talks with a sample of 25 7-11 year-olds, different from those of the present or the first study. The derived questions concerned perceived pace and difficulty level of different lessons, individual reactions to having too high or too low a work load, coping or other strategies used by a child who sees he or she is not in step with the lesson, and emotional reactions at being ahead of or behind the rest of the class.

(ii) Using these questions, interviews were then carried out with all children, aged between 7 and 11, from four "directive" classrooms (N = 53). The children were interviewed individually and informally during normal classroom routine, over a four-week period. Children were already well familiar with the practice of one author (JI) going into a classroom and talking to them, and were assured that information would not be passed back to the class teacher.

Teacher interviews

The four classroom teachers, and a fifth teacher who was similarly directive, were asked questions relating to their general classroom management style and to their perception of the relationships between lessons as timetabled and the actual activities of the pupils in their classes. They were asked: Would you label your classroom as directive, autonomous or as somewhere between the two styles; would you say your paper timetable (shown to teacher) closely represents, moderately represents, or does not represent (a) the lessons and curriculum areas taught at the given times across the week; (b) the activities of pupils at these times; (c) pupils' classroom activities over the term to date; (d) pupils' classroom activities for the foreseeable future?

RESULTS

Children's data

Children's answers to interview questions provide the substance of the study, and are accordingly presented question by question.

(1) *Pace of curriculum*: "Do you find keeping up with your lessons very easy, very hard or just right?"

Surprisingly, all 53 children in each class spontaneously placed themselves in the "very hard" or "very easy" group, with none feeling the pace was "just right". Children who found the going very easy were dubbed "frontrunners", and their opposites who found the going very hard "backmarkers".

Table 2 shows the distribution of responses broken down on the basis of age group for these frontrunners and backmarkers. Although there is considerable variation between classes, it is fairly apparent that the majority of children (41 or 77 per cent) perceive themselves as "backmarkers", or generally behind in their lessons. Typical reasons offered for why the child always seemed to be behind were: "teacher... has rubbed it off the board before I can write it down/ explains things too fast/ just carries on talking without waiting". The remaining 23 per cent who saw themselves as generally being ahead of the rest of the class, the "front runners", offered predictable reasons such as "the work was too easy/ I'd already done it before/ I already knew how to do it".

TABLE 2

PREFERENCES FOR SELF- OR TEACHER-DETERMINED STRATEGIES FOR MAINTAINING CURRICULUM CONTACT, AS REPORTED BY BACKMARKERS AND FRONTRUNNERS

Class	Strategies of Backmarkers (N = 41)		Strategies of Frontrunners (N = 12)	
	Self Determined	Teacher Determined	Self Determined	Teacher Determined
10 - 11 yrs (N = 14)	8	0	0	6
9 - 10 yrs (N = 17)	17	0	0	0
8 - 9 yrs (N = 13)	9	0	0	4
7 - 8 yrs (N = 9)	7	0	1	1
Totals:	41	0	1	11

(2) *General strategies*: "Would you tell your teacher you were falling behind/getting too much in front, or would you just carry on?"

First, observations during validity sampling in Study One provide some useful context for this question. It had been noticed that children unable to finish set work before break would typically continue afterwards, despite teacher direction to the class as a whole to begin a different subject. If this tendency were general, weaker pupils would then of course experience fewer curricular units than pupils who kept up with the pace set by the teacher.

By far the commonest response to this question was, "just carry on", so that on the face of it a self-regulated rather than a teacher-regulated strategy was preferred. Table 2 shows the extent to which children's reported reactions were predominantly self- or teacher-regulated, as defined by the question. All 41 of the children who could not keep up (backmarkers) predominantly used some device of their own to handle the problem, whereas only one frontrunner did this. Conversely, nearly all the frontrunners used some teacher-regulated tactic. Therefore, it appears that in these four directive classrooms, where teachers professed a high degree of monitoring and control, the backmarker children who were never able to keep up with the curriculum pace for the most part took their own decisions on what to do about it.

(3) *Specific strategies*: "If you were doing maths/English and teacher said change to English/maths but you hadn't finished, what would you do?"

Table 3 shows the range of strategies used by both backmarkers and frontrunners. Note, first, that total strategies are some three times the number of children, because children typically declared several strategies; means ranged from 1.25 to 4.6 per child. Table 3 also approximately classifies each offered strategy as "passive" (P1 to P4) or "active" (A1 to A10). An obvious passive strategy, if strategy at all, would be simply to allow unfinished work to accumulate over the day, whereas an active strategy would be to take the work home.

TABLE 3

BREAKDOWN OF COPING STRATEGIES USED BY BACKMARKERS AND FRONTRUNNERS TO MAINTAIN CURRICULUM CONTACT

		Backmarkers (N = 41)	Frontrunners (N = 12)
P1	Allow backlog to cumulate over the day	41	
P2	Allow cumulation over several days	40	
P3	Just leave the work unfinished	11	
P4	Look at timetable/lang card/worksheet		5
A1	Try best to keep up generally	2	
A2	Increase speed of working	9	
A3	Ask for help from other children	15	
A4	Work during the free period	7	
A5	Work during play break	1	
A6	Take the work home	8	
A7	Do more of same activity		9
A8	Practise multiplication tables		6
A9	Do some reading		5
A10	Miscellaneous (e.g., do sums/write/Lego)		3

A indicates active strategy; P indicates passive strategy

As Table 3 indicates, backmarkers overall prefer passive rather than active strategies. Altogether, such passive strategies account for 92 or 68 per cent of the solutions offered by backmarkers. For example, every one of the 41 backmarkers allowed backlogs to accumulate over one day, and all but one reported allowing such accumulation over more than one day (see rows 1 and 2). Simply leaving the work was also a noticeable ploy (11). Active strategies included trying to speed up or just trying their best (11), carrying on working through the free period (7) or even taking work home (8). Working through playtime or lunchtime was never really a strategy, but this may be put down to the school rules concerning supervision and safety. In fact several backmarkers volunteered that they would have liked to take advantage of such breaks.

Table 3 shows that for frontrunners there were five broad strategies and that these were quite different from those of the backmarkers — perhaps inevitably since one group was trying to handle an increasing work load and the other to create a work load. The most common strategy for frontrunners was to continue doing more of the same (e.g., "I'd finished the classwork so my teacher said turn over and carry on with the next page."). The next most used strategy was practising tables and the third reading. These were followed by assorted strategies used by only one or two children. While only one of the frontrunners' strategy types is classified as passive, it is noteworthy that even the four active strategies have the appearance of time-fillers and seem to represent classroom management solutions rather than a creative use of spare time. This may reflect the fact that they were teacher-routed rather than self-regulated ways of coping (Table 2).

(4) *Affective consequences*: "How do you feel when you finish early/can't finish?"

Table 4 indicates the reported positive or negative feelings of frontrunners and backmarkers from being out of step with the taught lesson. The main contrast is in terms of negative feelings. While frontrunners report never having any negative feelings about finishing early, for the backmarkers negative feelings were the norm. The kinds of comments made were, "I'm useless/ I feel thick/ I don't feel good". The frontrunners, for whom the classroom was a place to succeed comfortably and excel, offered feelings which were expectedly positive. The kinds of comments now made were, "I feel good/ It's nice/ I feel I'm better than the rest".

TABLE 4

REPORTED FEELINGS OF BACKMARKERS (B) AND FRONTRUNNERS (F) CONSEQUENT ON MISMATCH WITH TEACHER-DIRECTED TIMETABLE

Class	Negative feelings about self		Positive feelings about self		No marked feelings	
	B	F	B	F	B	F
10 - 11 yrs (N = 14)	7	0	0	5	1*	1
9 - 10 yrs (N = 17)	13	0	3	0	1	0
8 - 9 yrs (N = 13)	9	0	0	4	0	0
7 - 8 yrs (N = 9)	5	0	0	2	2	0

* Child's comment projected blame on to teacher

Less predictably, it can be seen that three backmarkers in one class actually had positive feelings about their situation. This was evidenced by comments such as "I feel okay/ I feel good 'cos I do my best". Also, only one child out of the 53 even raised the possibility that it might be the teacher's fault that they were out of step: "He gives us work that's too hard".

DISCUSSION

The five teachers in Study Two took a global view of events, maintaining that the paper timetable was a pretty accurate representation of actual curricular exposure and fairly represented the amount of time each child would experience in the various curricular areas. They felt this cognitive framework was supported and not countermanded by experience in class to date and was further a reliable indicator of future experience patterns. Although the children in four of the classes saw themselves as never in step with the timetable, the teachers were not sensitive to anything other than sporadic and inevitable variations for particular children. These variations were believed to be self-correcting and were not seen as affecting the pattern of that child's general curricular experience, still less that of the class as a whole. Thus the teachers might be said to be operating at a more strategic level within which anomalies could be assimilated without compromising an essentially strategic disposition (cf. Galton, 1987a). Alternatively, it might be that teachers were being not so much strategic as teaching to the mean (Bennett *et al.*, 1984), so that neither the most nor the least able children were being properly engaged. Whatever the inference, the natural question to ask is whether the kind of "slippage" represented in Figure 1 really *matters*.

One kind of answer, from within Study One itself, is that it matters because there is no evidence of discrepancies being self-correcting: if the term had been longer one wonders where or even whether the discrepancies would eventually have bottomed out. The slippage also matters in the light of findings from Study Two where teachers seemed unaware that the majority of children were needing to find *general* adaptive strategies and not merely deal with an occasional catching up problem. In this enterprise, the majority of backmarkers were what was termed self-regulating (Table 2). This may appear as a kind of autonomy, but autonomy has a positive connotation which was not apparent in their preference for passive solutions, nor in the reported class feelings of these children (Table 4). Rather, it seems an autonomy by default, arising from children's being unwilling to advertise to the teacher yet again that they are failing. When backmarkers were later asked why they did not go to teacher for help with the work load, the most common replies related either to classroom management or to the teacher's response set: "You have to wait too long in the queue/ I/d always have to have my hand up and I'd look stupid", or, "She just tells you to get on with it/ He explains too quickly when you ask".

The most general-level question to ask is why should there be any other than trivial discrepancy between the teacher-assumed and the child-actual curriculum? The comments from teachers indicated that they were in some sense unaware of what was really happening in the classroom. That is, they had a perception of children if not doing the current lesson at least "about to start" or "having just finished" so that the misregistration between intended and actual would not be serious enough to alter their global view of the planned and actual curricula being highly congruent — for both the class as a whole and for individual children. Relevant here is the finding from the Ford Teaching Project that teachers saw classroom practice in line with their orienting aims rather than with their actual practice, so that their perceptions were often discrepant with those of the Ford researchers whose "aims" were of course different (Elliott, 1976). Comments one hears in staffrooms may also indicate considerable tolerance to slippage when external pressures are operative: "We never do much in the last week anyway", or "In the last weeks of term the kids are in and out of the classroom rehearsing this, that and everything."

Teachers of course monitor what their children are doing, yet observation suggests this monitoring can itself have a self-fulfilling quality about it (cf. Elliott, 1976). The teacher checks to see what the child is doing; if the child is doing something other than the current lesson, she asks why. some answer such as, "I'm just finishing/trying to finish", is then taken by the teacher as acceptable. In this way the child, and somewhat ironically for directive classrooms, the backmarker in particular, is determining his or her own classroom experience. It is as though the 15 or so child strategies in Table 3 lead to stylised transactions which become a kind of currency for dealings between child and teacher, and a currency whose value is understood by both. It is also an aspect of the hidden curriculum in that the process is not easily accessible to the child or teacher, and is most readily detected by the outside observer.

Evidently, the injunction to "first finish what you were doing" seemed to direct the children's efforts more effectively than did the classroom lesson at the time. The cumulative consequence of such a regime was that weaker children would each day develop an increasing backlog of work, thereby setting demands which could not be met for time reasons alone, even assuming the child were able enough. This backlog phenomenon does therefore seem again a significant part of the hidden curriculum, and it was clear from asking the children that negative self-concepts were being formed, and a sense of strain and in some cases even hopelessness comes out in their answers.

As the pace of the backmarker slows and the backlog increases, the number of effective periods taught as recorded by children — and confirmed by observer sampling — also becomes increasingly discrepant from the teacher's own record. If one were to extrapolate the kinds of individual difference shown in Table 1 over the period of a school year, some children would have only half the maths experience of some others. There is, then, the additional question of what this experience would amount to. Would such children really be doing maths or merely repeatedly confirming their own failure at the subject?

Only one of the 12 frontrunner children seemed to offer a strategy that was self-regulated — the remainder mainly or always looked for advice to teacher for what to do next. However, like the "autonomy" of backmarkers, such teacher dependence ought not too readily to be taken literally. One suspects it may be no more than a reflection of the good reciprocity that exists between higher ability children and their teachers such that any teacher-oriented initiative from the child will always anticipate a reaction that is confirmatory of success. In contrast, such talking to teacher will be avoided by backloggers as confirmatory of failure (cf. Worrall *et al.*, 1988).

A point of caution here is to be wary of any *necessary* association between frontrunner-backmarker status and actual ability. First, high ability children can become "bored backmarkers", and they may deal with the boredom or pointlessness of some essay, for example, by using bigger writing or leaving large gaps between words, as found by Galton (1987a). More generally, classroom-wise pupils may tacitly bargain with the teacher by offering good behaviour in return for undemanding work which they work at just fast enough to avoid being rewarded with another helping. Equally, a child perceiving him or herself as a frontrunner may not fit the observational facts either. In this study, for example, two children who reported themselves to be frontrunners were identified from classroom observation as really in the backmarker category. This apparent mismatch between ability and self-perception may be no more than overcompensation and therefore one of several kinds of coping device available to backmarkers.

The snapshots of pupil reactions and pupil types afforded by the present data are complementary to those emerging from the series of Oracle projects (Galton, 1987b). Children in the second Oracle study were able to express the disappointment felt when teacher made alterations to improve their work, their fear of querying what teacher was doing, and not knowing whether a question from teacher embodied a genuine concern about learning or was a coded message of some kind.

Whatever the argument as to reasons, the present two studies provide cases of traditional classrooms where there are substantial differences between (a) actual and believed curricular exposure (Study One); (b) child and teacher perceptions on pacing and the effective timetable (Study Two). For some of the 65 children, the backmarkers, the timetable is a tow rope dragging them along and always threatening to trip them, while for frontrunners the same rope is really a harness that holds them in check. Every child is either towed or restrained by the timetable; none seems to slot comfortably into it.

Could the discrepancies have risen artefactually? In Study One, could there have been problems of definition disagreements between teacher and children over what constituted, say, English, and what reading? This seems unlikely, since the children's data were in close agreement with those of the observer. Further, non-confusable subjects like music and art showed similar patterns. In Study Two, it is possible that the sharply dichotomous frontrunner-backmarker classification arose because some children detected a subtle cue in the question that they should declare themselves as one or the other. However this may be, the investigators had no awareness of any such cue, and certainly had no orienting hypothesis that such a dichotomy would be found.

Although the authors' own experience is that negotiating classrooms offer a better learning environment (Ingram and Worrall, 1987), the present account must not be taken to imply that directive environments are necessarily bad and that there is no more to be said. Rosenshine and Stevens (1986) in reviewing studies and personal reports find no evidence that "formal-directive" classrooms are undesirable places, remarking that, "teachers in formal classrooms are warm, concerned, flexible and allow freedom of movement. But they are also task oriented and determined that children shall learn" (p. 380).

Although on the face of it the child in the traditional-directive classroom ought to have little scope for constructing his or her own curriculum, in a curious way it turned out that this responsibility was thrust upon the child, if that child wished to maintain an appearance of contact with the classroom enterprise. It was a responsibility that any primary-school child might find daunting to carry. It would therefore become inevitable that substantial and, indeed, non-correcting discrepancies would arise between the curriculum being run at the front of the class and children's own experiences. Also inevitable would be the children's subsequent feelings, particularly the feelings of backmarkers, as they saw their attempts to keep in step being at best only partially successful.

Correspondence and requests for reprints should be addressed to John Ingram who is now at Christ Church College, Canterbury, Kent CT1 1QU.

REFERENCES

Bennett, N. (1976). *Teaching Styles and Pupil Progress*. London: Open Books.

Bennett, N., Desforges, C., Cockburn, A., and Wilkinson, B. (1984). *The Quality of Pupil Learning Experiences*. London: Erlbaum.

Deci, E., and Ryan, R. (1985) *Intrinsic Motivation and Self-Determination in Human Behaviour*. New York: Plenum

Desforges, C., and Cockburn, A. (1987). *Understanding the Mathematics Teacher*. Brighton: Falmer Press.

Doyle, W. (1986). Classroom organization and management. In Wittrock, M. (Ed.), *Handbook of Research on Teaching* (3rd ed.). New York: Macmillan.

Elliott, J. (1976). Preparing teachers for classroom accountability. *Education for Teaching*, 100, 49-71.

Farquhar, C., Blatchford, P., Plewis, I., and Tizard, B. (1987). Curriculum diversity in London infant schools. *Br. J educ. Psychol.*, 57, 151-165.

Galton, M. (1987a). Change and continuity in the primary school: the research evidence. *Oxford Rev. Educ.*, 13, 81-93.

Galton, M. (1987b). An Oracle chronicle: a decade of classroom research. In Delamont, S. (Ed.), *The Primary School Teacher*. Brighton: Falmer Press.

Ingram, J., and Worrall, N. (1987). The negotiating classroom: child self-determination in British primary schools *Early Child Dev Care*, 28, 401-415.

Rosenshine, B., and Stevens, R. (1986). Teaching functions. In Wittrock, M. (Ed.), *Handbook of Research on Teaching* (3rd ed.). New York: Macmillan.

Worrall, N, Worrall, C., and Meldrum, C. (1988). Children's reciprocation of teacher evaluations. *Br. J. educ. Psychol*, 58, 78-88.

(*Manuscript received* 21*st January*, 1988)

Br. J. educ. Psychol., 60, 63-75, 1990

WHEN PUPILS DO NOT UNDERSTAND THE DETERMINANTS OF THEIR SUCCESS AND FAILURE IN SCHOOL: RELATIONS BETWEEN INTERNAL, TEACHER AND UNKNOWN PERCEPTIONS OF CONTROL AND SCHOOL ACHIEVEMENT

BY RUTH BUTLER AND RONIT ORION
(School of Education, Hebrew University of Jerusalem, Israel)

SUMMARY: Both dispositional and attributional perspectives on perceptions of control assume that people have some idea of outcome determinants. This, however, may not always be so. This study tested the hypothesis that pupils do not always understand the determinants of their learning outcomes, and that such a sense of unknown control will be associated with poor achievement in school. The study was unusual in that it tapped both dispositional control perceptions, using the new MMPCC, and causal attributions for success and failure in a school examination. Subjects were 186 10-year-old Israeli pupils of heterogeneous SES. Results confirmed that unknown control emerged as a distinct dimension of perceived control in both dispositional and attributional measures, and was consistently associated with poor achievement in school. Attributions of test outcomes to unknown causes were strongly related to dispositional unknown control, and were not affected by success or failure. In contrast, dispositional and attributional internality were unrelated, and did not predict achievement. It is argued that internal beliefs reflect internalised social-educational norms, and as a result have less impact on motivation than perceptions of unknown control. The implications of these findings for educational practice and motivational interventions are discussed.

INTRODUCTION

MUCH recent research on achievement motivation and behaviour has focused on people's perceptions of the determinants of their outcomes. While the relevant literature is vast, one can identify two main approaches (c.f. Stipek and Weisz, 1981; Marsh *et al*, 1984). One, guided mainly by social learning theory and specifically by Rotter's (1966) concept of locus of control, has been concerned primarily with identifying individual differences or dispositional tendencies in perceptions of control, and with studying the relations between such tendencies and broad outcomes such as school achievement. The other characteristic of attribution theorists (Kelley, 1972; Weiner, 1985) has focused on identifying those situation-specific variables which produce reliable differences in causal perceptions across subjects, and with relating these differences to specific outcomes such as expectancy for, or affective responses to, success and failure (Weiner *et al.*, 1971, 1978).

These different foci have yielded important differences in the questions posed and the methodologies adopted to investigate them. However, the present study was designed to address two assumptions which are implicit in both traditions, but have not been systematically investigated by either. First, both approaches tend to assume that people have fairly clear perceptions of the determinants of their achievement outcomes, and that these determinants are well represented by the possibilities offered by the experimenter. Studies of spontaneous attributions (Bar-Tal and Darom, 1979; Elig and Frieze, 1979) have confirmed that the causes of ability, effort and task difficulty, which feature prominently in theoretical formulations and structured questionnaires, are also frequently cited by subjects asked to offer their own reasons for outcomes. However, these studies also revealed that subjects rarely offer luck or chance as a causal determinant of either success or failure.

In this case, one may wonder what people mean when they express agreement with chance or luck attributions on structured questionnaires. One possibility is suggested by Connell (1985), who also found that children interviewed about their successes and failures rarely used a chance attribution. However, responses related to "not knowing why" they experienced particular outcomes emerged quite frequently. Thus, agreeing that an outcome was due to luck may be another way of saying that one does not understand what caused it. Following this reasoning, Connell's (1985) Multidimensional Measure of Personal Control in Children (MMPCC) includes not only the customary Internal and External (Powerful Others) scales, but also a third, Unknown Control scale, which taps the degree to which respondents agree that they do not know the determinants of their outcomes. Connell found that high unknown control for academic outcomes among elementary school pupils was associated with low scores on standardised achievement tests, and on questionnaire measures of perceived competence, mastery motivation and autonomy. Moreover, scores on the Unknown Control scale were related more consistently to these variables than were scores on the Internal and Powerful Others scales. Indeed, Harter and Connell (1984) found that the best fitting model for explaining the relations between perceptions of control, motivational orientations and achievement was that which placed Unknown Control at the beginning of the causal chain.

These findings imply both that pupils do not always have clear ideas about the determinants of their achievement outcomes, and that such unknown control may play a significant role in achievement behaviour. On a theoretical level, these findings are relevant to the recent distinction drawn by Skinner and Chapman (1984) between perceptions of causality, or people's understanding of relations between causes and outcomes, and perceptions of control, or the degree to which they feel they can influence their outcomes. Skinner and Chapman suggest further that perceptions of causality should logically take precedence over perceptions of control, since one can only evaluate one's capacity to control one's outcomes if one has some idea how these are determined. However, it is not clear why some pupils profess not to understand the causes of their outcomes in the classroom. One possibility is that causal uncertainty may be high only when pupils are asked, as in the MMPCC, to relate to brief and general statements, which should provide fewer informational cues than concrete outcomes in the classroom. In this case, perceptions of unknown control may be of limited value in understanding attributional processes for successes and failures in the classroom.

However, several researchers have noted the paucity of specific information about the determinants of learning outcomes provided both by teachers' verbal messages (Blumenfeld *et al.*, 1983) and by the most common forms of evaluation — praise (Brophy, 1981) and normative grades (Butler and Nisan, 1986). Thus, some children, especially in the early grades, may indeed feel that they do not understand the reasons for their successes and failures in school. This may be especially marked among less successful pupils, whose negative outcomes may not be a function primarily of effort, the causal determinant valued and emphasised by pupils, teachers, and parents alike (Bar-Tal and Guttmann, 1981). Thus Butler (1986) interpreted her findings that the causal attributions of high SES pupils were similar, regardless whether they were asked to explain their own outcomes or those of high or low SES others, while the attributions of low SES pupils differed in each case, as an indication that only the former had developed a consistent, "general" model for explaining learning outcomes.

However, no studies to date have investigated the degree to which unknown control plays a role also in causal attributions for specific learning outcomes. Thus, the first aim of the present study was to tap pupils' perceptions of internal, teacher and unknown control with reference both to school outcomes in general and to their outcomes on a school examination. It was hypothesised that unknown control will emerge as a distinct causal dimension in both generalised and specific perceptions of control. In addition, it was hypothesised that low achievers will express higher unknown control than high achievers both for school outcomes in general, and for their outcomes on an examination. Finally, the above analysis suggests that unknown control should be negatively related to perceived internality, and especially to effort attributions.

Examining both generalised perceptions of control and attributions for a specific event enabled us to address a second assumption of both social learning (Rotter, 1975) and attribution (Weiner, 1985) traditions. Both perspectives assume that the individual's cumulative experiences in school shape his or her generalised perceptions of the determinants of outcomes, and these dispositional tendencies both influence his or her perceptions vis a vis each new outcome, and can be modified by new experiences. However, as Stipek and Weisz (1981) note, most studies of causal perceptions have tapped either generalised perceptions or specific attributions, but not both, and thus surprisingly little is known about the relations between children's general ideas about the determinants of success and failure in the classroom, and their explanations for actual learning outcomes.

This lacuna is problematic for both bodies of theory. It is difficult to apply the findings of attributional studies, which by and large are based on highly contrived experiments which employ novel tasks and ignore individual differences in achievement history and generalised perceptions of causality, to classroom settings. Here children face familiar activities, and, at least by fourth grade, have formed stable conceptions of their ability (Eshel and Klein, 1981), and possibly also of the causes of their outcomes (Stipek and Weisz, 1981). Similarly, the contributions of dispositional studies to understanding, for example, the relation between control beliefs and achievement, are limited as long as it is not clear that dispositional locus of control does indeed influence perceptions for actual learning outcomes in specific classroom situations.

The present study was expected to shed further light on the interplay of dispositional and specific perceptions of causality in two ways. First, it was hypothesised that dispositional tendencies to explain school outcomes in terms of internal, teacher or unknown factors will be related to perceptions of the determinants of specific outcomes on the examination. This prediction follows from both social learning and attributional approaches, and is important to validate both. A second investigation probed the impact of success and failure on both generalised perceptions and specific attributions, an issue treated differently by social learning and attribution theorists. Rotter (1975) argued that when people have considerable experience in a particular domain their generalised perceptions will be relatively unaffected by specific events, and will thus be similar for both success and failure outcomes. In contrast, attribution theorists (Weiner *et al.*, 1971; Weiner, 1985) view outcome as the major determinant of causal attributions, and studies with children confirm that they tend to accept more responsibility for success than for failure (Frieze and Weiner, 1971; Friend and Neale, 1972; Ames *et al.*, 1976; Butler, 1986).

However, as noted above, previous studies have not clarified whether these patterns hold also when individual differences in achievement history and dispositional preferences are controlled for. Moreover, the strong relations between unknown control and achievement suggest that this may distinguish even more clearly than the customary internal-external distinction between perceptions of causality for success and failure. In other words, children may be more likely to feel that they do not understand the causes of failure than of success. Indirect support for this suggestion is provided by findings that children attribute failure more than success to luck (Friend and Neale, 1972; Butler, 1986), if one can assume that luck attributions reflect unknown control. Here, it was hypothesised that pupils who think that they succeeded on the examination will be more likely to attribute their outcome to internal causes, and will be less likely to attribute it to teacher or unknown ones, than pupils who think they have failed, even after one controls for past achievement history and dispositional perceptions of causality. If confirmed, this prediction will also shed light on the development of perceptions of unknown control, and on the relation between such perceptions and achievement, by indicating that these are more likely to emerge in response to failure than to success.

To conclude, the present study was designed to test three main hypotheses. First, it was hypothesised that unknown perceptions of control will emerge as a distinct dimension of pupils' perceptions of causality for school outcomes in general and for outcomes on an examination in particular, and that both dispositional and specific unknown control will be negatively related to school achievement. Second, it was hypothesised that dispositional perceptions of control will be correlated with corresponding dimensions of causal attributions for a specific learning outcome. Finally, it was hypothesised that pupils will express higher internality and greater understanding of the causes of outcomes for success than for failure, even after controlling for individual differences in achievement and dispositional control preferences.

Data relevant to testing these hypotheses were gathered in two stages. First, the Multidimensional Measure of Personal Control for Children was administered to 4th grade Israeli pupils. Second, about four weeks later, pupils attributed their outcomes on a school examination to internal, teacher or unknown causes corresponding to those tapped by the MMPCC.

METHOD

Sample

The MMPCC was administered to 237 4th grade Jewish Israeli pupils (107 girls and 130 boys) attending two elementary schools in a major Israeli city. Both schools are defined as integrated (i.e. with a mix of about 60 per cent middle-to-high SES and about 40 per cent low SES pupils) by the Israeli Ministry of Education. While 207 pupils from the original sample were also present when the examination was taken and when it was returned, data were analysed for 186 (104 boys and 82 girls), since in one class the examinations were returned without grades.

Measures

(1) General perceptions of control. The Multidimensional Measure of Personal Control for Children (Connell, 1985) was used. The measure has acceptable reliability and has been extensively validated (Connell, 1985). It consists of 48 items which tap Internal, Powerful Others and Unknown Control perceptions in four achievement domains: scholastic, social, physical activity and general. Six of the 12 items for each domain relate to success outcomes and six to failure ones. Respondents rate their agreement with each item on a four-point scale, where 4 indicates high and 1 low agreement. The measure was independently translated into Hebrew by two bilingual researchers. The two versions were very similar and minor differences were easily resolved. Data were analysed only for the 12 items in the scholastic domain.

(2) Attributions for test grades. Pupils first indicated whether they had succeeded, not succeeded, or did not know how they had done on the examination. They then rated the degree to which they agreed that the following causes affected the grade they had received.

Internal:	My ability in this subject.
	The effort I put into preparing for the examination.
Powerful Others:	How the teacher feels about me.
	The way the teacher taught the material.

In addition they rated their agreement with two unknown control statements:

I don't know what my grade depends on.

There's no way of knowing what determines grades on this examination.

Items were designed to correspond to the equivalent scales of the MMPCC, and their order was randomised across respondents. Pupils rated their agreement on four-point scales, with higher scores reflecting higher agreement.

(3) School achievement. Children's teachers rated each child's school grades and academic competence on two scales anchored at 1 (very poor) to 7 (outstanding). Since these measures were highly correlated, r (186) = 0.81, they were combined to form a single score. This index was standardised (X = 0, SD = 1.0) for each class.

(4) Examination grades. All teachers gave numerical grades out of 100. These too were converted into standardised scores as described above.

Procedure

The MMPCC was administered in each class by a female graduate student, following Connell's procedure. Questionnaires were identified by number, not name, and it was emphasised that responses were for research and would not be shown to teachers. Some two weeks later, pupils took an examination in mathematics, about which they were informed in advance, as part of their regular study programme. About ten days later, the experimenter entered the classroom immediately after the graded examinations had been returned. She explained that she was interested in pupils' thoughts about the determinants of their outcomes on the examination they had taken. The Attributions for Test Grades questionnaire was then distributed. Pupils indicated whether they had succeeded or not succeeded and then rated their agreement with the six statements described above. After completing the questionnaire, pupils were engaged in a discussion about the study.

RESULTS

Preliminary analyses yielded no significant differences in the results for boys and girls, so data for the sexes were combined.

Dimensions of control

General perceptions of control. Means and standard deviations for the Internal, Teacher and Unknown Control scales of the MMPCC, are presented in Table 1. The results show the tendency noted by Connell (1985) for elementary school children to endorse internal control items more than teacher or unknown ones. Following Connell's procedure, responses to all 12 items were analysed using a principal components factor analysis with oblique rotation. The analysis yielded three factors with eigenvalue > 1, broadly similar to those reported by Connell. The first factor accounted for 22.7 per cent of the variance and included all four unknown control items, together with one teacher success item ("The best way for me to get good grades is to get the teacher to like me"). The second factor accounted for 15.1 per cent of the variance and included all four internal items. The third factor accounted for 11.9 per cent of the variance and included three of the four teacher control items. Thus, with the exception of one item the results replicated the three dimensional structure hypothesised and reported by Connell.

Attribution for test grades. Means and standard deviations for each item and for the internal, teacher and unknown control composites are presented in Table 1. The results indicate that pupils continued to endorse internal more than teacher or unknown items, as in the MMPCC. Since this questionnaire included only two items for each dimension of control, correlations between all six were calculated, in preference to factor analysis. As expected, the correlations between the two internal items, r (186) = 0.24, and between the two unknown ones, r (186) = 0.45 were significant ($P<0.001$).

The teacher items were not correlated with one another. Instead, the one relating to the teacher's teaching was significantly correlated with the internal items, r (186) = 0.35 with ability and r (186) = 0.26 with effort, and the one relating to the teacher's attitude was correlated, but moderately, ($P>0.01$) with one of the unknown ones. One can note that this lack of differentiation was somewhat similar to that received for the teacher scale of the MMPCC. However, with the exception of the teacher items, causes were not significantly correlated across dimensions. Thus, the results as a whole seem to confirm that both questionnaires elicited broadly similar response tendencies, and that generalised and specific perceptions of causality for school outcomes seem to share a common structure. They also confirm that pupils endorsed unknown control items for specific no less than for general achievement outcomes (see Table 1).

TABLE 1

MEANS AND STANDARD DEVIATIONS FOR AGREEMENT WITH EACH SCALE OF THE MMPCC AND WITH ITEMS AND SCALES OF THE ATTRIBUTIONS FOR TEST OUTCOME QUESTIONNAIRE

	M	SD
MMPCC		
Internal Control		
All Items	3.70	0.48
Teacher Control		
All Items	1.61	0.64
Unknown Control		
All Items	1.88	0.88
Attributions for Test		
Ability	3.06	1.00
Effort	2.83	1.11
Internal Attributions	2.94	0.83
Teacher's Attitude to Me	2.53	1.11
Way Teacher Teaches	1.18	0.54
Teacher Attributions	1.85	0.64
I don't know	1.97	1.04
There's no way of knowing	2.04	1.05
Unknown Attributions	2.00	0.89

Relations between general and specific perceptions of control

Table 2 presents the intercorrelations between general and specific perceptions of internal, teacher and unknown control. In view of the non-significant correlation between the two teacher causes of test outcomes reported above, intercorrelations were also calculated for each item separately. The hypothesis that dispositional perceptions of control, as measured by the MMPCC, will be positively correlated with causal attributions for specific learning outcomes was not supported for internal perceptions. Generalised teacher control was moderately correlated with attributions to the teacher's attitude, but not to the teacher's teaching. In contrast, endorsement of unknown control items on the MMPCC was highly correlated with agreement that pupils did not understand why they had received the grades they did on the test. In addition, Table 2 indicates some relation between perceptions of unknown and of teacher control. Essentially identical patterns of intercorrelations were received also when these were computed separately for high and low achievers, defined as pupils scoring at or above the median, or below the median on teacher-rated achievement. The only clear difference was in the relation between internal and unknown beliefs. Among high achievers, the correlation was negative for both dispositional, $r\,(98) = -0.21$, and specific measures, $r\,(98) = -0.19$, $P<0.01$, while these dimensions were not correlated among low achievers.

Relations between perceptions of control and achievement

Correlations between generalised and specific perceptions of control, and the two measures of achievement are presented in Table 3.

TABLE 2

INTERCORRELATIONS FOR GENERAL AND SPECIFIC PERCEPTIONS OF UNKNOWN, INTERNAL AND TEACHER CONTROL

	1	2	3	4	5
General					
1. Unknown					
2. Internal	-0.05				
3. Teacher	0.36***	-0.08			
Specific					
4. Unknown	0.43***	-0.10	0.09		
5. Internal	0.07	-0.08	-0.11	0.03	
6. Teacher	0.23***	-0.06	0.13*	0.15*	0.36***
7. T's Attitude	0.28***	-0.11	0.24**	0.14*	0.08
8. T's Teaching	0.10	-0.02	0.03	0.10	0.38***

* P < 0.05; *** P < 0.001

Note: Correlations between general and specific agreement within each dimension of control are underlined. Intercorrelations between specific teacher attributions reported in text.

Again, correlations for the two teacher items of the Attributions for Test Grades questionnaire are reported separately. Table 3 indicates, somewhat surprisingly, that internal perceptions as tapped by both the MMPCC and the specific attributions questionnaire were not related either to school achievement in general, or to scores on the examination. As expected, endorsement of teacher items on the MMPCC was negatively correlated with both achievement measures, but these correlations were low. Attributions to the teacher's attitude were negatively correlated with general achievement. In contrast, unknown control as tapped by both questionnaires was significantly and negatively related to both measures of achievement. Thus the results indicate that perceptions of unknown control were more powerfully and consistently related to school achievement than teacher perceptions, and that internal perceptions were not related to achievement. In addition, one can note the basically similar patterns of correlations with achievement yielded by both general and specific measures of perceived control.

TABLE 3

CORRELATIONS BETWEEN PERCEPTIONS OF CONTROL AND SCHOOL ACHIEVEMENT

	Teacher Ratings	Grade on Test
General Perceptions of Control		
Internal	0.10	0.03
Teacher	-0.17**	-0.20**
Unknown	-0.28***	-0.26***
Attributions for Test Grades		
Internal	-0.09	-0.02
Teacher	-0.12	0.05
T's Attitude	-0.12*	-0.03
T's Teaching	-0.05	0.10
Unknown	-0.26***	-0.27***

* P < 0.05; ** P < 0.01; *** P < 0.001

Impact of success and failure on perceptions of control

General perceptions — MMPCC

Means for agreement of high and low achievers with success versus failure items within each control dimension of the MMPCC are presented in Table 4. The impact of outcome on pupils' endorsement of each dimension of control was analysed using 2 (success vs. failure) x 2 (high vs. low achievers) ANOVAs with outcome as a within and achievement as a between subjects factor. The analysis for internal perceptions of control yielded only a significant main effect for outcome, F (3, 182) = 8.34, $P<0.01$. As hypothesised, at both levels of achievement children agreed more with internal causes for success than for failure outcomes. The analysis for teacher perceptions yielded a significant main effect for achievement, F (3, 182) = 5.17, $P<0.05$, which reflects the previously noted tendency for low achievers to express higher teacher control than high achievers. The effect for outcome was only marginally significant, F (3, 182) = 3.29, $P<0.07$, but was in the expected direction. As hypothesised, the analysis for unknown control yielded significant main effects for both achievement, F (3, 182) = 16.64, $P<0.001$ and outcome, F (3, 182) = 30.95, $P<0.001$. Pupils at both levels of achievement endorsed unknown items more for failure than for success, and low achievers expressed higher unknown control than high achievers for both success and failure. Thus, the results supported the hypothesis that pupils will express greater understanding of the determinants of success than of failure, and will accept more responsibility for success.

Attributions for test outcome

Means and standard deviations for three levels of perceived outcome among high and low achievers are presented in Table 4.

TABLE 4

MEANS AND STANDARD DEVIATIONS FOR EACH DIMENSION OF GENERAL AND SPECIFIC CONTROL, BY OUTCOME AND SCHOOL ACHIEVEMENT

		Success		Failure		Unknown Outcome	
		High	Low	High	Low	High	Low
General Control							
Internal	M	3.78	3.75	3.61	3.65		
	SD	0.48	0.50	0.70	0.59		
Teacher	M	1.43	1.71	1.63	1.77		
	SD	0.65	0.86	0.78	0.93		
Unknown	M	1.48	1.96	1.94	2.28		
	SD	0.69	0.95	0.80	0.88		
	N	108	78				
Specific Control							
Ability	M	2.88	3.52	2.65	2.71	2.10	2.25
	SD	1.11	0.74	1.11	1.07	1.20	1.14
Effort	M	3.25	3.31	2.61	3.17	2.50	2.67
	SD	0.83	1.00	1.02	0.98	1.35	1.07
Internal	M	3.07	3.41	2.63	2.94	2.30	2.45
	SD	0.74	0.61	0.78	0.84	1.11	0.89
T's Method	M	2.91	2.89	1.90	2.23	1.70	2.58
	SD	0.92	1.08	1.04	1.17	0.95	1.16
T's Attitude	M	1.10	1.34	1.26	1.17	1.03	1.17
	SD	0.43	0.77	0.68	0.38	0.25	0.58
Unknown	M	1.81	2.14	1.95	2.27	1.60	2.55
	SD	0.85	0.87	0.93	0.86	0.87	0.83
	N	68	32	30	34	10	12

Internal causal attributions. In view of the distinction drawn by attribution theorists between internal attributions to ability, a stable internal cause, and effort, an unstable one, responses were analysed for the internal composite and for each internal determinant separately. Agreement with the internal composite for test outcomes was analysed with a 3 (perceived outcome: success-failure-don't know) x 2 (high vs. low school achievement) ANOVA. The analysis yielded a significant main effect for perceived outcome, F (3, 182) = 13.10, P<0.001. Duncan's test for a posteriori comparisons between means, using harmonic means to correct for unequal cell sizes, confirmed that agreement with internal causes was significantly higher (P<0.05) after perceived success (M = 3.17) than after perceived failure (M = 2.78) and significantly higher after perceived failure than among children who did not know whether they had succeeded or failed (M = 2.39). In addition, one can note a marginally significant tendency for low achievers to endorse internal perceptions more strongly than high achievers, F (3, 182) = 3.70, P<0.10.

The analysis for attributions to ability yielded similar results. At both levels of achievement, pupils who thought that they had succeeded endorsed ability more than ones who thought that they had failed, with ones who could not evaluate their outcomes agreeing least with this determinant, F(2.183) = 8.84, P<0.001. There was also a significant effect for achievement, F(2,182) = 4.67, P<0.05, with low achievers attributing outcomes, and in particular perceived success, *more* to ability than high achievers. The analysis for attributions to effort yielded a significant main effect for perceived outcome, F(3,182) = 6.73, P<0.01. At both levels of achievement, pupils attributed perceived success more to effort than perceived failure, with pupils who could not evaluate their outcomes endorsing this cause least. It is noteworthy that here too there was a marginally significant tendency for low achievers to endorse effort attributions more than high achievers, F(3,182) = 3.28, P<0.07, especially if they thought that they had failed.

Teacher attributions. In view of the poor correlation between the two teacher items, agreement with each was analysed separately. The analysis for attributions of outcome to the way the teacher taught the material yielded a significant effect for perceived outcome, F(3,182) = 15.26, P<0.001. Contrary to hypothesis, Duncan's test indicated that children who thought they had succeeded expressed significantly *higher* agreement with this cause (M = 2.90) than ones who thought that they had failed (M = 2.07) or ones who could not evaluate their outcome (M = 2.16). The analysis for attributions to the teacher's attitude to the pupil yielded no significant main or interaction effects.

Unknown attributions. In view of the strong correlation between general and specific perceptions of unknown control, the impact of perceived outcome on agreement with unknown items was analysed with a 3 x 2 ANCOVA, with endorsement of unknown items of the MMPCC as the covariant. The analysis yielded a significant main effect for achievement, F(4, 181) = 8.60, P<0.01, which reflects the previously noted negative correlation between unknown control for test outcomes and teacher ratings of achievement. The hypothesis that pupils would express greater unknown control for failure than for success was not supported; however, one can note the strong endorsement of unknown control among low achievers who could not evaluate their outcomes (see Table 4).

In summary, these results indicated that perceived success was associated with higher attributions to effort, ability and teaching method than perceived failure among both high and low achievers. In contrast, attribution of test outcomes to unknown causes was associated primarily with low school achievement and with generalised perceptions that one does not understand the causes of learning outcomes.

DISCUSSION

This study addressed two main questions: the role of unknown control in children's perceptions of the determinants of learning outcomes, and the relation between generalised and specific perceptions of control. While the results shed light on both questions, the most striking findings indicated that high and low achievers differed consistently in both general and specific perceptions of unknown control, while displaying intriguing similarities in their perceptions of internal control. Discussion will focus primarily on these results, since they have important implications not only for theory and research on perceptions of control but also for educational practice and intervention.

The strong negative associations between school achievement and agreement that one does not understand the determinants of learning outcomes in the classroom confirm and expand on findings from other studies (Harter and Connell, 1984; Connell, 1985). The factor structure, and mean agreement with each dimension of perceived control, were very similar to those reported for North American elementary school samples (see Connell, 1985), and thus provide preliminary confirmation of the cross-cultural utility of the MMPCC and of the unknown control construct.

In addition, the present results indicated that such general perceptions seem to affect causal understanding for concrete outcomes too. Thus, pupils who professed not to understand the determinants of learning outcomes in general were more likely to agree that they did not understand the determinants of their grade on the test. Moreover, low achievers were also more likely than high achievers to agree that they did not know why they received the grades they did. Taken together with the further finding that unknown control was more powerfully and consistently associated with achievement than either internal or teacher perceptions, these results support theoretical analyses which distinguish between understanding why outcomes occur and beliefs about one's control over one's outcomes (Skinner and Chapman, 1984; Skinner *et al.*, 1988) and suggest that the former may play a particularly important role in promoting competent behaviour.

Further research is necessary to clarify whether high and low achievers in this study differed in their sense of unknown control because the latter have not as yet worked out what outcomes in class depend on, or because low achievers have actually learned that there are no clear reasons for their own particular outcomes, as theories of learned helplessness suggest (Abramson *et al.*, 1978). However, the findings for internal control may shed some light on this question.

In contrast with perceptions of unknown control, perceptions of internal control did not differentiate between high and low achievers, since all pupils tended to agree that both school outcomes in general and outcomes on the examination depended primarily on internal causes. There are some indications that teachers' verbal messages about the causes of learning outcomes emphasise effort, and barely mention other causes (Blumenfeld *et al.*, 1983), and that internal attributions, and in particular effort, are favoured not only by the students themselves (Butler, 1986; Skinner *et al.*, 1988), but also by parents and teachers (Bar-Tal and Guttmann, 1981). The implication that beliefs that one's outcomes depend, or should depend, on oneself, reflect a pervasive social-educational norm is further supported by evidence that internal beliefs are perceived as highly socially desirable (Stern and Manifold, 1977; Weary *et al.*, 1985).

In this case, it seems that both high and low achievers in the present sample had worked out what outcomes in school should depend on, and thus that the latter too had a fairly clear idea of the attributional norms characteristic of classroom settings. While one can not generalise too far from the results of a single study, other studies using the MMPCC also report inconsistent relations between internal perceptions of control and school achievement (Harter and Connell, 1984; Connell, 1985; David and Connell, 1985). In addition, examination of the correlations reported by Stipek and Weisz (1981), in their extensive literature review, between achievement and internality as tapped by a variety of measures reveals a total of 105 significant correlations, as compared with 106 non-significant ones. Internal control beliefs as tapped by the most recently developed measure for children, the Control, Agency, and Means-end Interview are also not related to school achievement (Skinner *et al.*, 1988).

The rarer studies of the relation between achievement and attributions of specific outcomes to internal causes have also yielded mixed findings. Thus Stipek and Hoffman (1980) found that low achievers were more likely to attribute failure to lack of ability than high achievers, while Butler (1986) found no differences in their attributions of failure to ability, as in this study.

This apparently tenuous relation between internality and achievement makes sense if acceptance of responsibility for outcomes reflects internalisation of a social norm, rather than beliefs about the actual causes of one's own outcomes. In this case, the finding that general and specific perceptions of internal control were not related is not surprising. While few studies have examined the relation between dispositional perceptions of control and causal attributions for specific outcomes, some other investigators have also found that subjects who scored high on a dispositional measure of internal locus of control were no more likely than externals to endorse internal attributions for a specific outcome (Satterly and Joan, 1983; Lamb *et al.*, 1985).

At this point, one can note that pupils at all levels of achievement in the present study also agreed that the way the teacher taught the material was an important determinant of test grades, and that attributions to this cause were highly correlated with attributions to both effort and ability. This finding seems strange, in view of the clear distinction between internal and external sources of control drawn by both social learning (Rotter, 1966, 1975) and attribution (Weiner *et al.*, 1971; Weiner, 1985) theorists. However, another study of attributions for natural achievement outcomes in the classroom also found that the way teachers teach was associated with internal, rather than external, items and was highly endorsed by pupils, teachers and parents alike (Bar-Tal and Guttman, 1981). Similarly, Skinner *et al.* (1988) report that pupils at ages 7-12 were more likely to agree that outcomes depend on teachers' explanations than on their helpfulness or liking for the pupil, and that the former did not load on the same factor as the latter items.

If internality and attributions to the teachers' teaching reflect acceptance of social norms, one can understand why endorsement of these determinants did not distinguish between high and low achievers. Thus, at least in the present sample, beliefs about causality in the classroom seemed to fall into two distinct clusters. One, which is shared by most pupils, maintains that achievement in general is the responsibility of the pupil and that test outcomes depend on effort, ability and the way they are taught. The other, which clearly distinguished between low and high achievers, maintains that one does not understand the causes of learning outcomes. While endorsement of internal beliefs was associated with low unknown control among high achievers, low achievers seem to hold two apparently contradictory perceptions. On the one hand they accept responsibility for their outcomes, but on the other they do not understand how these are determined.

One possibility is that agreement with unknown items reflects an ego-protective strategy, whereby it is easier for failing students to say that they do not know why they failed than to attribute negative outcomes to their own lack of ability or effort. This is consistent with the finding that pupils expressed higher unknown control for failure than for success on the MMPCC, and expressed higher internal control for success. However, this discrepancy was more marked for high than for low achievers. Moreover, perceived failure on the test was not associated with higher agreement with unknown items and low achievers were even more likely to attribute failure to effort, ability and the teacher than were high achievers. Thus, taken together, the results support the interpretation that successful and unsuccessful pupils differed more in the degree to which they understood the causes of their outcomes than in their acceptance of responsibility for them.

If valid, this analysis of the similarities and differences in the control perceptions of high and low achievers has important implications for both theory and practice. On a theoretical level, the present findings seem to confirm the importance of distinguishing between perceptions of causality and of control, and suggest that dispositional internal control may not be as useful as is commonly believed in characterising the causal beliefs of high and low achievers. In addition, the results underline the importance of unknown control in characterising control beliefs at different levels of school achievement, and suggest that by fourth grade uncertainty about the determinants of learning outcomes in general represents a stable and salient attitude which shapes attributions for specific outcomes too, and is not modified by specific success or failure outcomes.

On an applied level, these findings suggest first that teachers should incorporate explicit and ongoing information about the reasons for learning outcomes into their verbal communications and evaluations of pupils' work. While this recommendation may seem rather obvious, Blumenfeld *et al.* (1983) found that less than 1 per cent of teachers' communications related to the determinants of pupils' outcomes. In this case, it is hardly surprising that many pupils do not understand why they succeed or fail. Similarly, asking pupils if they understand why they attained one outcome rather than another should promote causal understanding and constructive attributions (Nisan, 1981) in high and low achievers alike, and could help teachers identify, and plan interventions for, those children high in unknown control. Second, these findings have implications for designing such motivational interventions. The strong endorsement of internal causes by both high and low achievers suggests that the most common strategies, whereby pupils are encouraged to adopt internal perceptions of control (Reimanis, 1974; Wang and Stiles, 1976) or to attribute failure to lack of effort (Dweck, 1975), may not be very effective in classrooms where most pupils already accept personal responsibility for their outcomes. Indeed, such beliefs may even contribute to less able pupils' sense of unknown control, if, for example, they accept that failure is due to insufficient effort, but see that they do poorly even when they try hard.

While further research is necessary to clarify the determinants of unknown control, especially in low achievers, the present results suggest two possible avenues for intervention. One could focus specifically on pupils' sense of unknown control, by encouraging them to explore why they do not understand the determinants of learning outcomes. Enhanced understanding of why and when they do not understand could then help pupils determine both when they can affect their outcomes, and how they might do so. Another might focus more generally on the relation between attributional norms and educational outcomes. Even though teachers convey to pupils that their outcomes depend on effort, they usually evaluate achievement in terms of outcome relative to others (Levine, 1983). This double message may in itself contribute to the unknown control of low achievers, and this should be alleviated if teachers were more consistent, and evaluated pupils' work in terms of individual mastery, progress and effort, rather than relative outcomes and achievement. In addition, recent theory and research on motivational orientations suggests that such a shift from ability-focused, outcome criteria to individual, learning ones will enhance continuing motivation (Butler, 1987; Dweck and Leggett, 1988), and achievement (Nicholls, 1984) among high and low achievers alike.

Correspondence and requests for reprints should be addressed to Dr R. Butler, School of Education, Hebrew University of Jerusalem, Mt. Scopus, Jerusalem, Israel 91905.

REFERENCES

Abramson, L. Y., Seligman, M. E. P , and Teasdale, J. (1978). Learned helplessness in humans: a critique and reformulation. *J. abnorm Psychol.*, 87, 49-74.

Ames, C., Ames, R., and Felker, D. W. (1976). Informational and dispositional determinants of children's achievement attributions. *J. educ Psychol* , 68, 63-69.

Bar-Tal, D., and Darom, E. (1979). Pupils' attributions for success and failure. *Child Dev.*, 50, 264-267.

Bar-Tal, D., and Guttmann, J. (1981). A comparison of teachers', pupils', and parents' attributions regarding pupils' academic achievement. *Br J. educ. Psychol.*, 51, 301-311.

Blumenfeld, P. C., Hamilton, V. L., Bossert, S. T., Wessels, K., and Meece, J. (1983). Teacher talk and student thought: socialization into the student role. In Levine, J. M., and Wang, M. C., (Eds.), *Teacher and Student Perceptions: Implications for Learning*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Brophy, J. E. (1981). Teacher praise: a functional analysis. *Rev. educ. Res.*, 51, 5-32.

Butler, R. (1986). The role of generalised expectancies in determining causal attributions for success and failure in two social classes. *Br. J. educ. Psychol.*, 56, 51-63.

Butler, R. (1987). Task-involving and ego-involving properties of evaluation: the effects of different feedback conditions on motivational perceptions, interest, and performance. *J. educ. Psychol.*, 79, 474-482.

Butler, R., and Nisan, M. (1986). Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *J. educ. Psychol.*, 78, 210-216.

Connell, J. P. (1985). A new multidimensional measure of children's perceptions of control. *Child Dev.*, 56, 1081-1041.

David, H. B., and Connell, J. P. (1985). The effect of aptitude and achievement status on the self system. *Gifted Child Q.*, 29, 131-136.

Dweck, C. S. (1975). The role of expectations and attributions in the alleviation of learned helplessness. *J. Person. soc. Psychol.*, 31, 674-685.

Dweck, C. S., and Leggett, E. (1988). A social-cognitive approach to motivation and personality. *Psychol. Rev.*, 95, 256-273.

Elig. T. W., and Frieze, I. H. (1979). Measuring causal attributions for success and failure. *J. Person. soc. Psychol.*, 37, 621-634.

Eshel, Y., and Klein, Z. (1981). Development of academic self-concept of lower-class and middle-class primary school children. *J. educ. Psychol.*, 73, 287-293.

Friend, R., and Neale, J. (1972). Children's perceptions of success and failure: an attributional analysis of the effects of race and social class. *Dev. Psychol.*, 7, 124-128.

Frieze, I. H., and Weiner, B. (1971). Cue utilization and attributional judgements for success and failure. *J. Person. soc. Psychol.*, 39, 591-606.

Harter, S., and Connell, J. P. (1984). A model of the relationships among children's academic achievements and their self-perceptions of competence, control, and motivational orientation. In Nicholls, J. (Ed.), *The Development of Achievement Motivation.* Greenwich, CT.: JAI.

Kelley, H. H. (1972). *Attribution in Social Interaction.* New York: General Learning Press.

Lamb, R., Lalljee, M., and Jaspars, J. (1985). A paradoxical prediction from locus of control. *Person. indiv Diff.*, 6, 277-291.

Levine, J. M. (1983). Social comparison and education. In Levine, J. M., and Wang, M. C. (Eds.), *Teacher and Student Perceptions: Implications for Learning.* Hillsdale, N.J.: Erlbaum.

Marsh, H. W., Cairns, L., Relich, J., Barnes, J., and Debus, R. L. (1984). The relationship between dimensions of self attribution and dimensions of self concept. *J. educ. Psychol.*, 76, 3-32.

Nicholls, J. M. (1984). Achievement motivation: conceptions of ability, subjective experience, task choice and performance. *Psychol. Rev.*, 91, 328-346.

Nisan, M. (1981). Motivational aspects of evaluation in school In Lewy, I. A. and Nevo, D. (Eds.), *Evaluation Roles in Education.* Hillsdale, N.J.: Erlbaum.

Riemanis, S. (1974). Effects of locus of reinforcement control modification procedures in early grades and college students. *J. educ. Res.*, 68, 124-127.

Rotter, J. B. (1966). Generalised expectancies for internal versus external control of reinforcement. *Psychol. Monogr.*, (Whole No. 609).

Rotter, J. B. (1975). Some problems and misconceptions related to the construct of internal versus external control of reinforcement. *J. consult clin. Psychol.*, 43, 56-67.

Satterly, D., and Joan, H. (1983). Personality differences and the effects of success and failure on causal attributions and expectancies of primary school children. *Educ. Psychol.*, 3, 245-248.

Skinner, E. A., and Chapman, M. (1984). Control beliefs in an action perspective. *Hum. Dev.*, 27, 129-132.

Skinner, E. A., Chapman, M., and Baltes, P. B., (1988). Control, means-ends, and agency beliefs: a new conceptualization and its measurement during childhood. *J. Person. soc Psychol.*, 54, 117-133.

Stern, G. S., and Manifold, B. (1977). Internal locus of control as a value. *J. Res. Person* , 11, 237-242.

Stipek, D., and Hoffman, J. (1980). Development of children's performance-related judgements. *Child Dev.*, 51, 912-914.

Stipek, D. J., and Weisz, J. R. (1981). Perceived personal control and academic achievement. *Rev. educ. Res.*, 51, 101-137.

Wang, M. C., and Stiles, B. (1976). An investigation of children's concept of responsibility for their school learning. *Am. educ. Res. J.*, 13, 159-179.

Weary, G., Jordan, J. S., and Hill, M. G. (1985). The attributional norm of internality and depressive sensitivity to social information. *J. Person. soc. Psychol.*, 49, 1283-1293.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psych. Rev.*, 92, 548-573.

Weiner, B., Frieze, I. H., Kukla, A., Reed, L., Rest, S., and Rosenbaum, R. M. (1971). *Perceiving the Causes of Success and Failure.* New York: General Learning Press.

Weiner, B., Russell, D., and Lerman, D. (1978). Affective consequences of causal ascriptions. In Harvey, J. H., Ickes, W. J., and Kidd, R. F. (Eds.), *New Directions in Attribution Research (Vol. 2).* Hillsdale, N.J.: Erlbaum.

(*Manuscript received* 3rd *January*, 1989)

Br. J. educ. Psychol., 60, 76-87, 1990

READING ILLUSTRATED SCIENCE TEXTS: A MICRO-COMPUTER BASED INVESTIGATION OF CHILDREN'S STRATEGIES

By D. J. REID and M. BEVERIDGE
(Centre for Science and Technology Education, University of Manchester)

SUMMARY. Research has shown that the effect of illustrations on children's learning of science is related to ability. This study examines the hypothesis that children who show different levels of success in a learning task will employ different strategies in their picture-text processing. One hundred and eighty 14 year-old children were required to learn three illustrated science topics of different levels of difficulty which were presented to them on a BBC micro-computer. The computer was programmed to record the time spent on each sentence and each picture, as well as the point in the text at which the picture was accessed. Post hoc testing enabled the children to be divided into six groups according to their success rates on the learning tasks. Significantly longer times were spent looking at the pictures as the topics increased in difficulty; also the amount of time spent looking at pictures increased as the learning of the children decreased. For every second the least successful children spent accessing pictures, they spent about four seconds reading the text; for the most successful children this ratio was about one-to-six. The least successful children also accessed the pictures significantly more frequently than the most successful. The different strategies used by the children are discussed in terms of what is known about their differential learning gains from illustrated texts.

INTRODUCTION

IN recent years research has indicated some of the conditions under which a positive effect of pictures on children's learning of science might be expected to occur. For example, a positive picture effect is more likely if the information to be learned is present in both picture and text (Levie and Lentz, 1982), and also when learning outcomes are specified in terms of recognition memory rather than a form of "understanding" that requires the learner to infer information beyond the given facts (Donald, 1979; Reid *et al.*, 1983). We have also been able to demonstrate a differential picture effect amongst the least and most able secondary school children (Reid and Beveridge, 1986). In the latter study the most able children performed almost 20 per cent better when pictures were present, whereas the least able children remembered significantly less when pictures were present. A positive correlation between reading ability and the effect of pictures has also been found for primary school children, using miscue analysis (Beveridge and Griffiths, 1987). These results suggest that children of different abilities might use different strategies when learning from illustrated text. The study reported in this paper investigates the way children of different abilities use pictures as adjuncts to text processing.

Recent work using micro-computers has shown how real-time reading strategies might be investigated. It is possible to present text in a number of different ways on a computer driven VDU. Each mode of presentation provides a somewhat different analog of the unconstrained reading process which operates when text appears on paper. Aaronson (1984) and Aaronson and Ferres (1984) used a word-by-word self-paced presentation of sentences, each new word extinguishing the previous one. This allows the precise time spent on each word to be recorded. However, in order to make the task more like normal reading, Mitchell and Green (1978) enabled children to move on by presenting words in groups. Even this task was more constrained than in normal reading since each new group of words appearing on the VDU replaced the previous group; but, according to the authors, it "encourages readers to keep up to date with their processing". Other work (Beveridge and Edmundson, 1989) has modified this technique to allow children to regress to earlier parts of the text.

The advantages that these techniques have is that they allow accurate measures of time and order of processing for each unit of text. However, no work has been carried out using similar techniques to study the effect of pictures on text processing. The study reported here presents science texts together with pictures which can be accessed on the screen by using a computer key.

In the present study new text did not replace old text. Once units of text (entire sentences) had appeared on the VDU they were allowed to remain. This was because when the children called up the picture on the screen, it replaced the text. It was judged that children, when returning to the text from the picture, should have all the text that they had read available to them, as would be the case in normal reading.

The aims of this study are, then, (a) to investigate the picture-text processing strategies used by a cohort of children learning from illustrated science texts, and (b) to relate these strategies to the success (assessed by a criterion-referenced test) which the children showed on the learning tasks.

METHOD

Materials

Three biological texts were written for use with 14-year-old children. These were "Transport of sugar in plants", "Exchange of gases in the leaf", and "Conduction of the nerve impulse". Each was explanatory rather than narrative in style, and contained between 232 and 245 words in either eleven or twelve sentences. They were written at increasing levels of difficulty based on three criteria: independent measures of readability (Reid, 1984), an analysis of the linguistic structure of the texts (Perera, 1987), and the conceptual level of the science content using Shayer and Adey's (1981) classification schema. A single picture was produced with each topic and drawn in colour with an AMX mouse, for presentation on a BBC micro-computer. Each picture encapsulated the main conceptual contents of its associated text in a form familiar to the children (*).

The computer was programmed to allow the reader to access either a picture or a new sentence of text by pressing one of two keys. Having viewed the picture, the reader could also access directly the text previously presented. The program also recorded the time spent viewing the picture and text on each occasion they were accessed. The point in the text at which the picture was called up was also recorded. This technique, therefore, allowed for the recording of:

(i) time on reading each sentence in the topic,

(ii) time on reading the picture,

(iii) frequency of picture access, and

(iv) position of picture access in relation to the latest sentence called up.

Three criterion-referenced tests were developed along the lines reported in earlier studies (Reid *et al.*, 1983; Reid and Beveridge, 1986). They were specially written such that the information required for the correct answer was available in both picture and text. Examples of test items are given in Table 1. The children were required to complete one short six-item test upon completion of each program before they were allowed to proceed to the next program. The total score (/18) made by the children on these tests provided the data base for the post hoc rank ordering of the subjects into six groups from those learning least to those learning most (N = 30 for each group). As a check that the topics increased in difficulty in the direction predicted by the writing criteria, a two-way correlated ANOVA was also performed on the results. This indicated highly significant differences between the mean scores for the three tests (F = 81.827, df = 2:358, P<0.01). Post hoc t-testing indicated that all of the means were significantly different from each other in the predicted direction.

(P<0.01 in all cases.) The Tukey conservative two-tailed t-test is used in all analyses of differences between means in this report.

TABLE 1

EXAMPLES OF TEST ITEMS

Text	Objective item	Relevant quote
Easy	The ringing experiment shows that A) *sugar is transported in the phloem* B) water is not transported in the phloem C) water travels up the plant D) the leaves produce sugar	This shows that *phloem transports sugars* from the leaves down the stems
	Sugar cannot get across the 'ring' because A) the plant is dying B) *the phloem tubes have been cut* C) there is not enough water to dissolve it D) the leaves have stopped producing it	If a horizontal ring of phloem is removed from a stem with a sharp knife, then the *sugars from the leaves cannot bridge the gap across the cut tubes*
Medium	Transport of gases in the leaf occurs A) through the water supply B) via the leaf cells C) *through the stomata* C) via the leaf epidermis	These gases enter the plant through small holes in the underside of its leaves called *stomata*
	Transport of oxygen in or out of the leaf occurs A) when carbon dioxide transport has stopped B) only in the light C) *at the same time as that of carbon dioxide* D) is dependent on the carbon dioxide concentration	So, by diffusion, the *transport of oxygen and carbon dioxide occurs* in opposite directions, but *at the same time*
Difficult	When an axon is not conducting it A) has too few sodium ions B) has too few potassium ions C) *is negative on the inside and positive on the outside* D) is positive on the inside and negative on the outside	The membrane surrounding the resting (i.e. non-conducting) neuron is polarised, with the *inside negatively charged relative to the outside*
	The Na+/K+ exchange requires A) electricity B) high concentration of sodium C) high concentration of potassium D) *energy*	The NA+/K+ exchange requires *energy* which is supplied by ATP

Note: Recognition phrases are italicised here, but were not italicised in the original text or test

Sample

Participants were 180 14-year-old children, comprising the complete second year of a large co-educational city comprehensive school, and hence a typical range of abilities. None of the children had studied the topics previously.

Procedures

The children worked in groups of 10 in the school computer laboratory. An additional introductory program was used to familiarise the children with the user-friendly instructions that appeared on the screen, and three adults were present to answer any questions about the procedure. The orders of presentation of the topics was distributed equally between the children. The exercise was self-paced, with no child taking longer than one hour to complete.

The children were told that they would be given a brief test on how well they had learned the material at the end of the session. As already indicated, the children called up a topic one sentence at a time by pressing the 'W' (writing) key, but sentences already on the screen were not replaced in the process. In this way the complete topic was built up. At any point between sentences the children could call up the picture associated with the text by pressing the 'P' (picture) key. This replaced the text with the full-screen picture. When the 'W' key was pressed again, the picture was replaced by the screen of text previously available plus the next sentence in the sequence.

RESULTS

The results are presented in two sections. In the first section the time spent on text and picture in all three topics together is reported. In the second section the detailed results for one of the topics (Topic 1) are presented. An analysis of the way the children used the pictures in association with the initial sentence in each of Topics 1 and 2 is also given.

Time spent on text and pictures in all three topics

The times spent on the text and picture together in all three topics are given in Table 2. The means indicate a trend of decreasing time spent on the topics as the success of the children increases (correlation 0.884, $P<0.05$).

TABLE 2

TOTAL TIME IN SECONDS SPENT ON THE TOPICS

Variable		Mean	SD
Success group:	high 1	222.644	102.807
	2	263.356	117.983
	3	259.034	110.111
	4	262.066	115.492
	5	271.234	114.967
	low 6	291.156	118.669
Difficulty:	easy	240.056	107.179
	medium	267.756	113.033
	difficult	277.756	120.373
Mode:	text	219.469	93.606
	picture	42.133	38.326

Two-way ANOVAs were performed for the times spent separately on text and picture in the three topics. The results show that the text was looked at for significantly increasing lengths of time as the topics increased in difficulty, from an average of 205.706 seconds for the easiest text to 229.333 seconds for the most difficult text ($F = 4.776$, df 2:348, $P<0.01$). However, children in the different success groups did not spend significantly different total amounts of time on the texts of the three topics taken together ($F = 1.461$, df 5: 174, NS), neither was there a significant interaction between difficulty of text and success of child on text reading time ($F = 0.562$, df 10: 348, NS).

The results for time spent on pictures showed that these were also looked at for significantly longer periods of time (34.350 seconds to 48.200 seconds) as the topics became more difficult ($F = 8.268$, df 2:348, $P<0.01$). But, unlike the time spent accessing the text, the time spent on pictures did relate significantly to the success of the child ($F = 2.761$, df 5: 174, $P<0.05$). Thus, the amount of time spent looking at the pictures significantly decreases as the success of the child increases (see column (b) Table 3, and Figure 1). There is, however, no significant interaction between topic difficulty and success of child on picture viewing time ($F = 0.371$, df 10: 348, NS).

TABLE 3

RATIOS OF MEAN TOTAL TIME ON TEXT AND PICTURE FOR THE DIFFERENT SUCCESS GROUPS (IN SECONDS)

Success group		combined time on all texts (a)	combined time on all pictures (b)	ratio (a/b)
High	1	191.267	31.389	6.090
	2	226.644	37.933	5.969
	3	218.178	40.522	5.384
	4	217.711	44.333	4.911
	5	228.272	43.011	5.306
Low	6	235.656	55.500	4.200

FIGURE 1

TIME SPENT LOOKING AT PICTURE AS A FEATURE OF CHILD SUCCESS

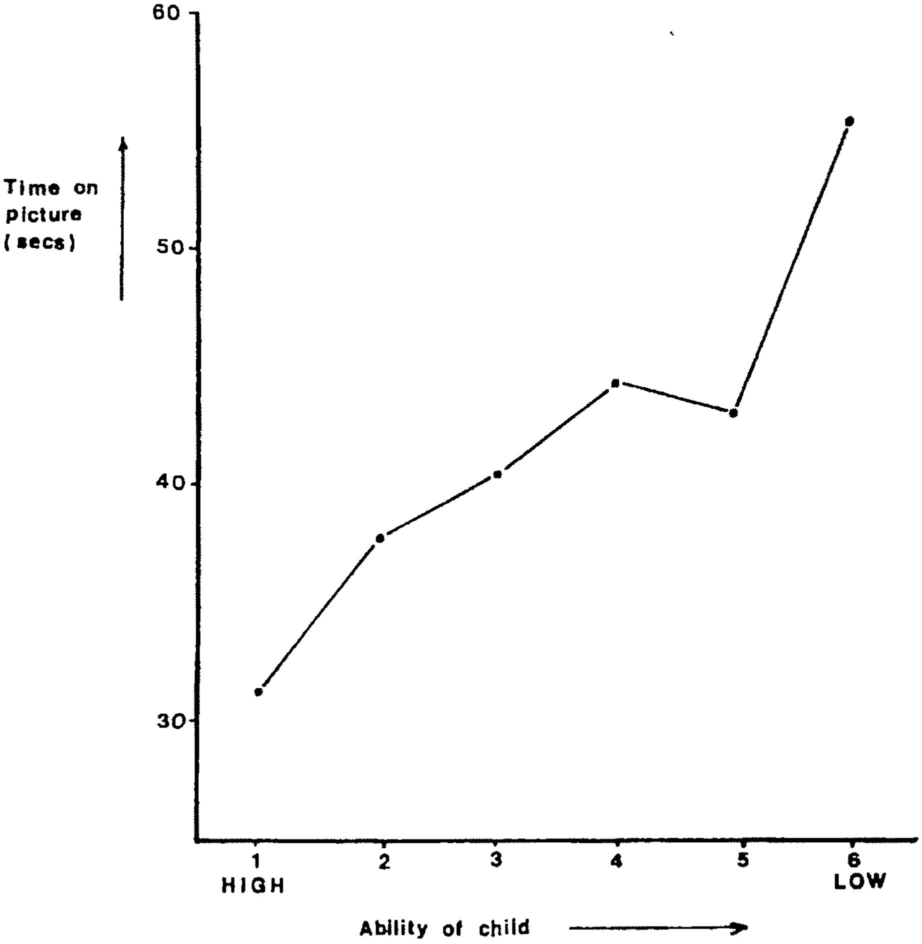


Table 3 also shows that the ratio of time on text: time on picture varies across the success range. There is a significant correlation between these two variables, the ratio of text: picture time decreasing as the success of the children decreases (regression coefficient = –0.9116; P=0.013). For every 6 seconds the most successful children spend reading the text they spend 1 second accessing the picture; this contrasts with a ratio of about 4 to 1 for the least successful children.

Text and picture selection strategies for Topic 1

Topic 1 is the easiest of the topics by all four of the criteria used to describe it. However, it is characteristic of a type of written science material which 14-year-old children commonly use in schools in the U.K.

Strategies can be considered from two points of view. They can be considered in terms of the frequency and the points in the text at which the children access the pictures. Alternatively, the influence of the pictures on the time taken to process the text can be considered.

The total number of times the picture was accessed in Topics 1 and 2 is shown in Table 4. The trend in frequency of picture access for Topic 1 was measured using Page's L test (Cohen and Holliday, 1982). The test is significant (L = 1006.5; C = 6; S = 13; P<0.05) and the calculated regression slope is shown in Figure 2 (correlation between frequency of access and success = 0.826; F = 8.606, df 1: 4, P<0.05; slope = 3.429 ± 3.246).

TABLE 4

TOTAL NUMBER OF TIMES THE PICTURES ARE ACCESSED IN TOPICS 1 AND 2

Success group		No. of times the pictures are accessed	
		Topic 1	Topic 2
High	1	22	31
	2	34	49
	3	39	44
	4	34	47
	5	44	49
Low	6	41	58

Clearly, the less successful children access the picture significantly more frequently than the more successful. This suggests that the earlier result which indicated that less able children spend proportionally more time looking at the pictures (Table 3) is associated with calling up pictures more frequently. An identical pattern is found for Topic 2 (L = 978; C = 6; S = 12; P<0.01, and the correlation between frequency of access and ability in this case = 0.834; F = 9.124, df 1: 4, P<0.05; slope 3.943 ± 3.626).

Attention is now turned to the amount of time children spend reading the different sentences in the text. When they first start to read a topic, some children choose to look at the picture before they look at any of the text, others go directly to the text itself. Comparison of the viewing times for the first sentence of the topic under these two picture conditions can reveal the influence of picture on individual sentence processing time. Table 5 shows the differences in reading time for the first sentence in Topics 1 and 2 under two different picture conditions. When the picture is accessed prior to the sentence being viewed (the PIC-SEN condition) both sentences are viewed for a relatively long time. In this condition there is no significant difference between the mean times spent on each sentence. However, when the picture is not viewed before the sentence is seen (the O-SEN condition), there is a significant difference in the time spent reading the two initial sentences. The most likely explanation for this comes from the considerable difference in the amount of new information in the two sentences which, in turn, is reflected in their grammatical complexity. The first sentence of Topic 1 is "Some plants are quite tall". The first sentence of Topic 2 is "In much the same way as each individual cell in an organism requires its own supply of food and water, so it also requires its supply of oxygen and, in the case of plants, of carbon dioxide". The interesting, and surprising, result shown in Table 5 is that when a picture is viewed prior to reading these sentences the effect on reading time of sentence complexity is not found. The prior viewing of the picture increases the reading time for both sentences and eliminates the difference between them.

FIGURE 2
RELATION BETWEEN FREQUENCY OF PICTURE ACCESS AND THE SUCCESS OF THE CHILD

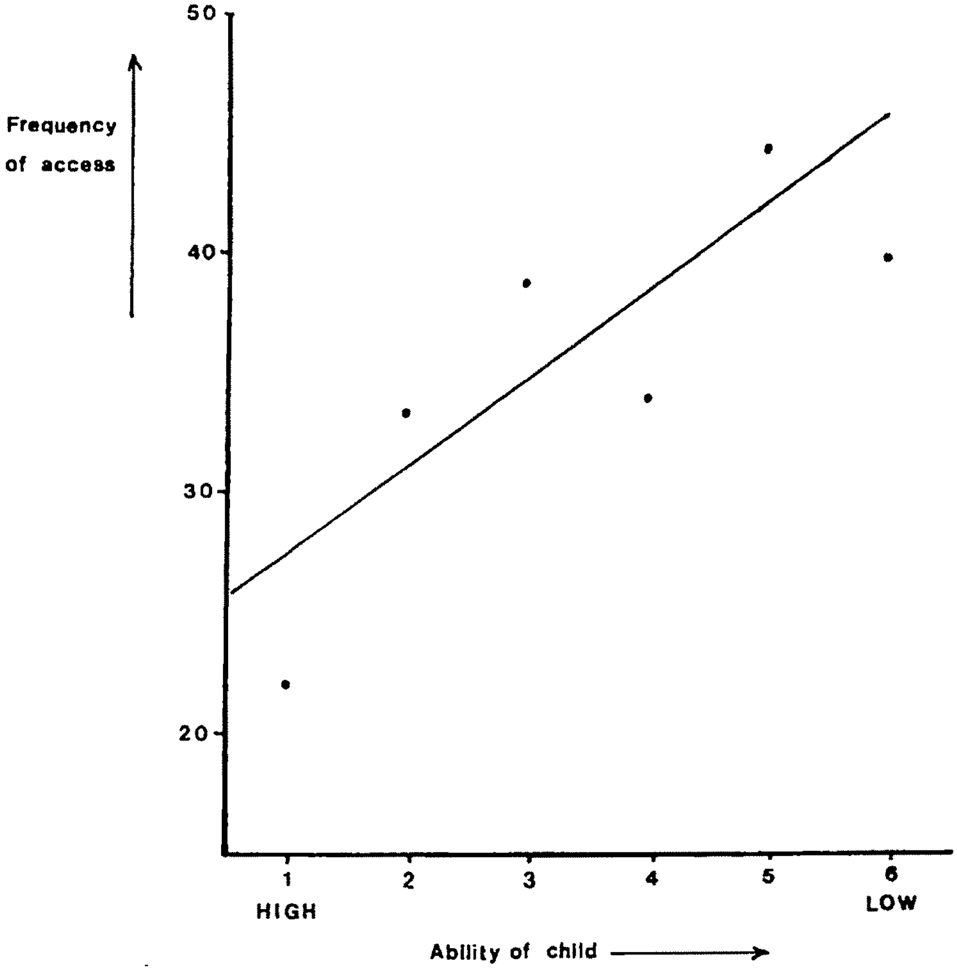


TABLE 5

MEAN TIME (IN SECONDS) CHILDREN SPEND ACCESSING THE FIRST SENTENCES IN TOPICS 1 AND 2 UNDER TWO PICTURE CONDITIONS

Picture condition	Time in seconds		t-test for independent means
	Topic 1	Topic 2	
PIC-SEN	32.59 (N = 111)	28.97 (N = 111)	1.58 (NS) (df = 220)
O-SEN	4.49 (N = 69)	17.18 (N = 69)	7.96 (P < 0.01) (df = 136)

Note: It is coincidental that the number of children accessing the picture prior to the sentence is the same for both topics. They were not always the same children,

This influence of picture on text processing found for the initial sentences can be further studied for the whole text. Three rankings of sentence viewing were obtained, based on the total amount of time the 180 children spent viewing each of the sentences in Topic 1 under three different conditions of picture-text relationship, namely:

(i) when the picture had been accessed immediately prior to the sentence being accessed (PIC-SEN condition),

(ii) when the picture had not been accessed immediately before the sentence was accessed (0-SEN condition), and

(iii) when the picture had been accessed immediately after the sentence had been accessed (SEN-PIC condition).

As suggested by the 0-SEN condition data for the first sentences of Topics 1 and 2, viewing time might also indicate the importance of a sentence to the information contained in the text as a whole. Independent evidence of the relevant importance of the sentences was obtained from two sources. Three expert judges, who were teachers familiar with the topic under discussion, were asked to rank the sentences in order of importance. These judges were in close agreement with each other (Spearman's rho = 0.544, 0.560 and 0.989; all significant at $P<0.01$). A second ordering of sentence importance was made by 183 schoolchildren in another school who were given the topic to read from a worksheet; their opinions were combined into an overall judgment, and the correlation between the teachers' judgment and the children's was 0.713 ($P<0.01$). Comparison of the relative viewing times for individual sentences under the three conditions with the expert judgments of importance will indicate whether the pictures were influencing text processing in the same way as for the initial sentences. Table 6 shows the correlations between ranked processing time for sentences under the three conditions and the order of the expert judgments of importance.

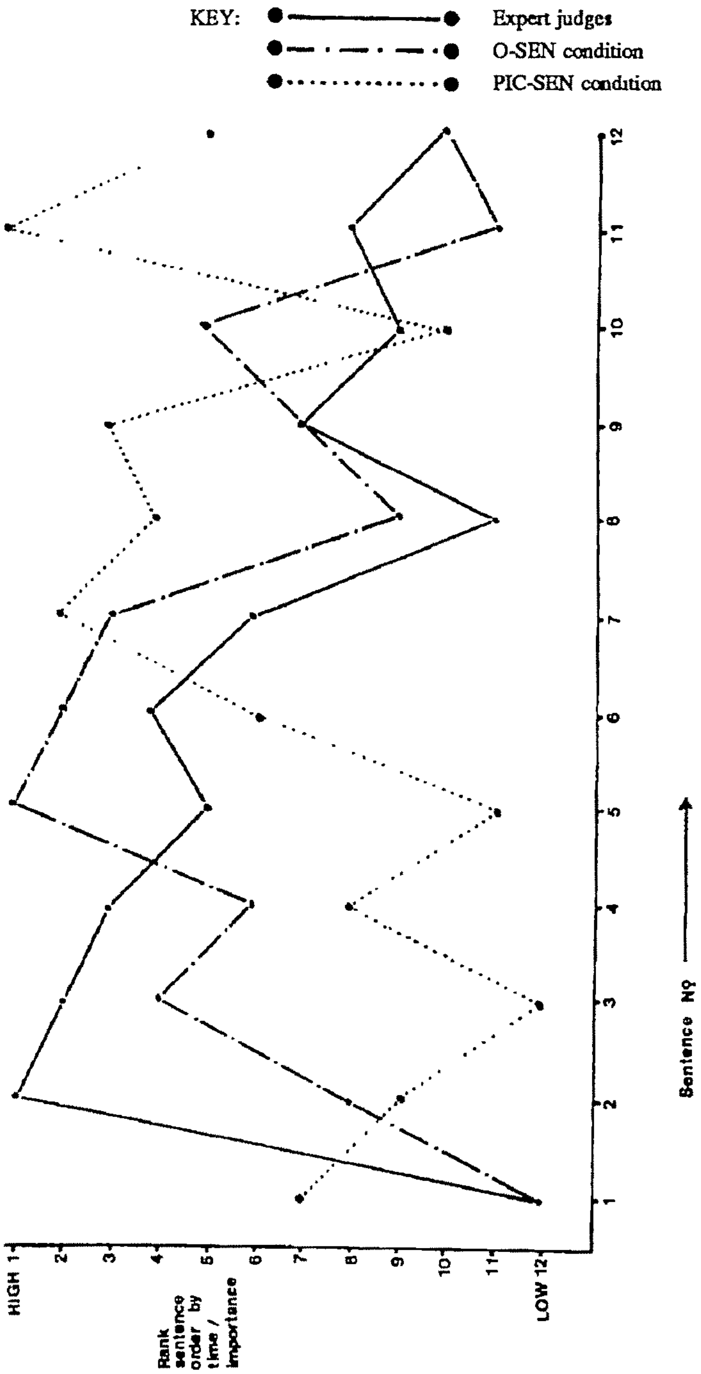
TABLE 6

SPEARMAN CORRELATION MATRIX OF SENTENCE RANKING BY TIME (UNDER THREE CONDITIONS) AND EXPERT JUDGES

Rankings based on:		1	2	3	4
PIC-SEN	1	–	-0.378	-0.028	-0.406
O-SEN	2	–	–	0.119	*0.580
SEN-PIC	3	–	–	–	0.007
expert judges	4	–	–	–	–

* $P < 0.05$

FIGURE 3
RANKING OF SENTENCES BY TWO CONDITIONS OF PICTURE VIEWING AND JUDGES' OPINIONS OF SENTENCE IMPORTANCE



Only one of the correlations in Table 6 is significant, that between the judges' opinion of sentence importance and the amount of time that the children spent viewing the sentences when they had not looked at the picture immediately prior to calling up the sentence on the screen. Perhaps equally as interesting is the negative correlation between the PIC-SEN condition and the expert judges. Once again, it appears as if prior viewing of the picture influences the amount of time spent reading the following sentence, and not necessarily in the direction which might be expected. We can show the detail of the processing patterns outlined in Table 6 in Figure 3. The almost parallel lines of the ranking of the expert judges and the rankings of the 0-SEN condition contrast sharply with the near mirror image lines of the PIC-SEN rankings. This result suggests, once again, that it is the use of picture which disrupts the relationship between sentence importance and sentence processing time. Only when the picture is not accessed does the time spent reading a sentence correlate significantly with its agreed importance in the text.

DISCUSSION

The research reported here takes as its point of departure the previously reported result that the addition of pictures to a science text correlates significantly with improved learning by the most able children and with decreased learning by the least able children (Reid and Beveridge, 1986). Self-paced presentation of text and pictures using a computer driven VDU allows investigation of the strategies used by children with different learning successes.

It has often been demonstrated that poorer readers take longer to read non-illustrated material (e.g., Aaronson and Scarborough, 1976). The present study finds a similar significant trend for less successful learners who spend longer reading illustrated topics. In fact, they spend up to twice as long looking at pictures as the more successful (Figure 1).

One purpose of this study was to provide information as to why pictures differentially influence learning according to children's abilities, a result first identified by Reid and Beveridge (1986). Since, in the present study, the children who learn least from the text spend marginally longer looking at it (Table 3) we can reject a reduction in text processing time in the presence of pictures as a simple explanation of their poorer performance on illustrated texts. The improved performance of more successful children on illustrated texts suggests that they are able to identify and integrate relevant information from both texts and pictures. They can look to the pictures for clarification and are not distracted from the text by them. This suggests that the points in the text at which pictures are accessed, together with the frequency of access, may be associated with children's ability. In the present study the analysis of text and picture selection strategies for Topic 1 investigated these questions. (The results for Topic 2 replicated the results for Topic 1, although for brevity they are, for the most part, not included in this report. Unfortunately the children do not learn enough from Topic 3 to allow a valid detailed interpretation of those results.)

From the analysis of Topic 1, it is clear that less successful children looked at pictures both more frequently and for a greater proportion of the time they spent on the topic. As a consequence, they face the task of integrating pictorial and textual information to a greater extent than their more successful peers. Evidence that this integration process is important is shown by the fact that every time a picture was accessed, the succeeding sentence was looked at for a significantly longer time. We have already pointed out a methodological difficulty here. When returning to the text from the picture the reader has access to the entire text so far revealed. Thus, length of time alone does not reveal the precise cognitive activities which occur immediately after changing from picture to text. During this period the children will be engaged in the process of re-orientation, extraction and integration of information. It is not clear from the present data whether all or some of these processes take longer for the less successful children, or even whether all these children share similar processing strategies. Further research is required to distinguish between orientation, extraction and integration processes in text/picture switching.

Why do children access pictures? The data suggest that judgment of importance of a sentence does not predict the point at which pictures are accessed. It is possible that instructions given to the children at the outset of the exercise may have influenced their use of the pictures. It is difficult to explain to children how to call up a picture without alerting them to the fact that the pictures might play an important part in the topic. However, it has been noted that children find it difficult to ignore pictures, and that seems to be reflected in this study. In both Topics 1 and 2 almost two-thirds of the children elected to look at the picture even before they had called up the first sentence (111/180, Table 4. It is coincidental that the number is exactly the same for both topics — they were not all exactly the same children.)

The question as to whether less successful children access pictures at less appropriate points in their processing of texts is difficult to answer because of the number of different reasons why a picture could be regarded as helpful. Analysis of the data failed to demonstrate an interaction between location of picture access and subject success level, and further research is needed to investigate this possibility.

Biederman (1981) has demonstrated that meaning can be extracted from pictures "at a glance" (in as little as 37msecs). It is also known that human memory for pictures is remarkably good (Standing *et al.*, 1970). Whilst the pictures and information differ in these studies, it may also be that early exposure to a picture informs children sufficiently that subsequent accessing becomes less and less important, regardless of text content. There is some evidence for this interpretation from the data, since for Topics 1 and 2 both frequency and time of access to picture decrease the further into the text the children progress.

It is tempting to speculate that higher frequency of picture access by less successful children might reduce semantic cohesion between sentences (Halliday and Hasan, 1976). Inappropriate points of access and longer viewing times may interfere with memory for links in discourse. However, the effect may be in the opposite direction. Poor understanding of cohesive textual links may lead less successful children to look for help in the pictures. Again, research is required which can further relate text processing strategies to use of pictures.

Overall, this study indicates that there are systematic influences of pictures on text processing. However, it is probably unwise to continue to use the word "strategy" in this context of picture-text processing, implying as it does an element of planning and goals which the data have failed to uncover.

The effect of pictures on text processing has been shown to relate to the ability of the children to remember information in the topics. It is suggested that the earlier result (Reid and Beveridge, 1986), that pictures enhanced learning by more able children but decreased learning in less able children, is associated with these different processing patterns. This in itself suggests that learning might be improved if children were taught how to use pictures effectively.

*The authors are happy to provide black and white photocopies of these pictures and their associated texts to any interested teacher. The original pictures have been photographed in colour direct from the VDU screen, but, because of the colour, do not reproduce clearly enough to be worth including in this report.

Correspondence and requests for reprints should be addressed to Dr David Reid, Department of Education, University of Manchester, Oxford Road, Manchester M13 9PL.

REFERENCES

AARONSON, D. (1984). Computer methods and ecological validity in reading research. *Behav. Res. Meth , Instr. Comput.*, 16, 102-108.

AARONSON, D., AND FERRES, S. (1984). Reading strategies for children and adults: some empirical evidence. *J. verb. Learn. verb. Behav* , 9, 700-725.

AARONSON, D., AND SCARBOROUGH, H. S (1976). Performance theories for sentence coding: some quantitative evidence. *J. exp. Psychol.: Hum Percept.*, 2, 56-70.

BEVERIDGE, M., AND EDMUNDSON, S. (1989). Micro-computer based assessment of reading strategies of good and poor readers in word and phrase presentation. *J. Res. Reading*, 1-12

BEVERIDGE, M., AND GRIFFITHS, V. (1987). The effect of pictures on the reading processes of less able readers: a miscue analysis approach. *J Res. Reading*, 10, 29-42.

BIEDERMAN, J. (1981). On the semantics of a glance at a scene. In KUBORY, M. and POMERANTZ, J. R. (Eds.), *Perceptual Organization*. Hillsdale, N.J.: Erlbaum.

COHEN, L., AND HOLLIDAY, M. (1982). *Statistics for Social Scientists*. London: Harper and Row.

DONALD, D. R. (1979). Effects of illustrations on early oral reading accuracy, strategies and comprehension. *Br J educ. Psychol* , 49, 282-289.

HALLIDAY, M. A. K., AND HASAN, R. (1976). *Cohesion in English*, London: Longman.

LEVIE, W. H., AND LENTZ, R. (1982). Effects of text illustrations. *Educ. Commun. Technol.* 30, 195-232.

MITCHELL, D. C., AND GREEN, D. W. (1978). The effect of context and content on immediate processing in reading. *Q Jl. exp. Psychol* , 30, 609-636.

PERERA, K. (1987). Personal communication Department of Linguistics, University of Manchester.

REID, D. J. (1984). A three-in-one readability program for science worksheets. *School Sci. Rev.*, 65, 560-569.

REID, D. J., AND BEVERIDGE, M. (1986). Effects of text illustration on children's learning of a school science topic. *Br. J. educ Psychol.*, 56, 294-303.

REID, D. J., BRIGGS, N. AND BEVERIDGE, M. (1983) The effect of picture upon the readability of a school science topic. *Br. J educ. Psychol.*, 53, 327-335

SHAYER, M , AND ADEY, P. (1981). *Towards a Science of Science Teaching*. London: Heinemann.

STANDING, L , CONEZIO, J., AND HABER, R N. (1970). Perception and memory for pictures: single-trial learning of 2500 visual stimuli. *Psychon. Sci* , 19, 73-74.

(*Manuscript received 27th January*, 1989)

International Symposium
Research on

EFFECTIVE AND RESPONSIBLE TEACHING

September 3 through 7, 1990

in Fribourg, Switzerland

For information write to:

Walpen Marie-Tony, Congress Secretary
Paedagogisches Institut, University of Fribourg
Route des Fougères, CH-1700 Fribourg, Switzerland
Phone: (037) 219 630 FAX: (37) 219 650
E-Mail: dickand@cfruni53.bitnet

Br. J. educ. Psychol., 60, 88-92

TEACHER APPROVAL AND DISAPPROVAL IN HONGKONG SECONDARY SCHOOL CLASSROOMS

BY SAM WINTER
(Department of Education, Hongkong University)

SUMMARY. OPTIC (Observation of Pupils and Teachers in Classrooms) was used to study Hongkong secondary school teachers' use of approval and disapproval and pupils' on-task behaviour. Despite great individual differences, the sample of teachers was found overall to focus upon pupils' academic behaviour and to emphasise the use of approval. However, disapproval was emphasised as a means of responding to pupils' social behaviour. Positive correlations were found between approval (particularly to academic behaviour) and on-task behaviour. Negative correlations were found between disapproval (particularly to social behaviour) and on-task behaviour. In all these respects the findings were similar to those of a recent UK study.

INTRODUCTION

Classroom interaction research has a long history. Some of the more interesting research in the field over recent years has been in the area of teacher approval and disapproval and the relationship with pupil behaviour. These studies have spanned primary and secondary education sectors in a number of Western developed countries and have utilised a variety of observation instruments. In view of the different settings and instrumentation involved there is a strikingly consistent pattern in the data. That pattern may be summarised as follows: (1) teachers' general rate of responding (approval and disapproval combined) is lower with older age groups, (2) disapproval is often more frequent than approval, particularly at higher grades, (3) disapproval is particularly common in response to pupils' social behaviour. Some studies have attempted to identify pupil correlates of teacher approval and disapproval. For example, Nafpaktitis *et al.* (1985) found that rates of approval and disapproval were both correlated (negatively and positively respectively) with pupil off-task and disruptive behaviours. For a fuller review of this and other studies see Wheldall *et al.* (1989).

In the last few years some interesting work on approval, disapproval and pupil behaviour has been done using OPTIC (Observation of Pupils and Teachers in Classrooms), a schedule developed at the Centre for Child Study at the University of Birmingham, and described by Merrett and Wheldall (1986).

OPTIC provides a method for observing pupil behaviour and teacher responses in mainstream classrooms during periods in which pupils are engaged in activity individually, in groups or as a class. It is designed to be used over observation sessions of 30 minutes, half of which time focuses upon pupil on-task behaviour and upon teacher approval/disapproval respectively.

Teacher responses to pupil behaviour are recorded as positive (approving) or negative (disapproving), and are further categorised in terms of the type of pupil behaviour (academic or social) that has prompted the teacher's response. Additionally, responses may be categorised in terms of certain semantic features. Purely instructional teacher behaviour is ignored at all times during observation of teacher responses. Pupil behaviour is recorded by way of a simple system that enables each pupil's on-task behaviour to be observed individually. On-task behaviour is defined as eye-contact with the teacher or the task or otherwise following the teacher's instructions. For further details of OPTIC the reader is referred to Merrett and Wheldall (1986).

OPTIC data from secondary school classrooms

Wheldall *et al.* (1989) trained independent observers to employ OPTIC to record teacher and pupil behaviour in 130 secondary school classrooms in the UK. All classes were mixed ability. Pupils were aged 11 to 16. Most school subjects were observed. Mean class size was 22. Total observation time for each teacher was 45 minutes.

The data from the study by Wheldall *et al.* is summarised in Table 1. It shows that as a group teachers used approval more than disapproval, that they responded more to academic behaviour than to social behaviour, and that they employed disapproval more than approval when responding to social behaviour. Wheldall and his colleagues report that, in regard to social behaviour, women teachers seemed particularly prone to use disapproval. They also report that approval was given rather less where older classes were being taught. Their findings are therefore consistent with the pattern described in the opening paragraph of this paper.

TABLE 1

SUMMARY DATA FROM THE WHELDALL, HOUGHTON AND MERRETT (1989) STUDY: TEACHER RESPONSES

Pupil Behaviour Involved	Academic	Social	Total
Type of Response Given			
Approval	45	10	55
Disapproval	15	30	45
Total	60	40	100

Note: Figures indicate approximate percentages of total observations for the group concerned.

Wheldall *et al.* found that, over the 45 minutes of study, teachers averaged 52.9 responses of the type measured by OPTIC. However, there were great individual differences between teachers in their rate of responding: the standard deviation for this variable indicates that one in three teachers gave either less than 26 or more than 79 responses.

Wheldall *et al.* also report that the mean proportion of lesson-time spent on-task was 80 per cent. However, this figure again concealed a wide range, running from 17 per cent to 99 per cent in the classes observed. They add that they found a positive association between on-task behaviour and the use of approval to academic and to social behaviour, and the use of approval to academic and to social behaviour, and a negative association between on-task behaviour and disapproval to social behaviour. They stress that no causal connection can be inferred from this association.

This paper reports a Hongkong replication of the Wheldall *et al.* study. OPTIC was used to study patterns of teacher and pupil behaviour in Hongkong secondary school classrooms. Very little classroom interaction research has been conducted in Hongkong schools, and it was expected that distinctive characteristics of the education system (relating to organisational features, language of instruction, values and resources) might promote patterns of classroom interaction very different to those in the West. For further details of the Hongkong education system the reader is referred to Morris (1983, 1985, 1988), Johnson (1983), Hongkong Education Department (1986) and Hongkong Government (1987).

METHOD

Graduate students engaged in a postgraduate initial professional training course in education at Hongkong University and enrolled in a course on "Behavioural Approaches in the Classroom" were trained to use the observation schedule. The author first gave verbal instruction, then gave out written guidelines for using the observation schedule. The guidelines included samples of completed record forms. Then students practised using OPTIC in a video exercise which depicted a range of teacher and pupil behaviours in a role-play situation. Finally, students were given opportunities to ask questions related to the use of OPTIC.

Eighty-six students acted as observers in the present study. Forty students were studying on the full-time course. They were required to use OPTIC during a two-week period of school experience and were asked to observe a class taught either by a teacher employed at the school or by their student partner at the same school.

The remainder of the observers were drawn from a group of 80 part-time students. These students were all experienced teachers who, having obtained several years of teaching experience, were now intent on obtaining a professional qualification. These students were given the option of two exercises, one of which involved OPTIC. Forty six students from this group chose to do the OPTIC assignment. They were asked to observe classes taught by other teachers employed at the school at which they worked.

All observers were required to indicate basic details relating to the class being observed: the school, the form level, the number of pupils present and the status of the teacher being observed (student or staff). Date and time of the observation was noted. All observations were collected between November, 1986 and March, 1987.

Each observer observed one class for one 30-minute period. Class size averaged 39.13. The range was from 28 to 56 pupils. The classes represented Forms 1 (17 classes), 2 (22 classes), 3 (29 classes), 4 (13 classes) and 5 (5 classes). The majority of pupils observed would therefore be in the age range 12 to 17, although there would inevitably be a number of over-age and under-age pupils present in the sample.

In the 86 lessons observed, most subjects were represented. Fifty-four fell into the arts, social studies and language category, while 30 fell into the science and mathematics bracket. The nature of two was not recorded.

Observation time for this study was 43 hours; 21.5 hours for observation of teacher behaviour and an identical figure for pupil behaviour. During this period a total of 2033 teacher responses (approval and disapproval) and 19,300 pupil observations were made.

RESULTS

Separate analyses of observational data from the two groups of students revealed no substantial differences. No relationship was found between class size and on-task behaviour. On-task behaviour and teacher responses were found not to be associated with pupils' form level or type of subject taught ('arts' or 'sciences'). Consequently, analyses are henceforth presented for the entire sample of 86 teachers and classes.

On-task behaviour

Pupils were on-task for an average of 69.93 per cent of observations. The range for this variable was 25 per cent to 97 per cent, with a standard deviation of 16.46. These figures indicate that a significant number of classes in this study displayed rather low rates of on-task behaviour; in fact approximately 25 per cent of those studied displayed on-task rates under 60 per cent.

Teacher responses

Rates of teacher approval and disapproval to academic and social behaviour are displayed in Table 2. Responses to academic behaviour were overwhelmingly approving (approximately 50 per cent of total responses compared with 16 per cent for disapproval). In contrast, responses to social behaviour were more often disapproving (about 21 per cent of total responses) than approving (13 per cent). Responses to academic and social behaviour combined were overwhelmingly approving (around 63 per cent, largely because of the fact that most teacher responses were to academic behaviour (about 66 per cent).

TABLE 2

MEAN TEACHER RESPONSES FOR THE HONGKONG STUDY (SDs IN BRACKETS)

Pupil Behaviour Involved	Academic		Social		Total	
Type of Response Given						
Approval	50.38	(25.16)	13.02	(13.59)	63.40	(25.01)
Disapproval	15.94	(15.14)	20.66	(18.67)	36.60	(25.01)
Total	66.32	(20.78)	33.68	(20.79)	100.00	—

Note: Figures indicate percentage of total responses

As is evident from the standard deviations for these variables, these mean figures conceal massive differences in figures from individual teachers. Figures for academic approval ranged from 0 per cent to 100 per cent of total responses. The range for social approval was 0 per cent to 53 per cent. Those for academic and social disapproval were 0 per cent to 71 per cent and 0 per cent to 75 per cent respectively.

Although the general tendency was for teachers to give more approval than disapproval there were 22 teachers (27 per cent of the sample) whose disapproving responses exceeded their approval. Similarly, although the general trend was for teachers to emphasise academic behaviour, there were nevertheless 16 (19 per cent of the sample) who responded most frequently to social behaviour.

In terms of absolute numbers the mean total of responses given by teachers during the 15 minutes they were under study was 23.64 (SD 12.40). The range was again wide, one teacher failing to respond in any way during the period observed and one displaying 61 responses (about one every 15 seconds).

Correlations between different categories of teacher response

Teachers who emphasised approval also tended to focus upon pupils' academic behaviour. The Pearson's r correlation between these two variables (expressed in terms of proportion of total teacher responses) was 0.5387 (significant at the 0.001 level). Conversely, an emphasis upon disapproval appeared to be linked to a focus upon pupils' social behaviour. The correlation here was 0.5384 (again significant at the 0.001 level). All tests of significance were one-tailed.

Correlations between teacher responses and on-task behaviour

Strong associations were found between teachers' use of approval and disapproval (expressed as percentages of total teacher responses) and pupils' on-task behaviour. As may be seen from Table 3, higher rates of on-task behaviour were associated with greater emphasis upon approval and upon pupils' academic behaviour. Academic approval and social disapproval were both strongly correlated (one positively and the other negatively) with on-task behaviour. These figures should be interpreted in the light of the strong association, mentioned above, between approval and academic behaviour on one hand and between disapproval and social behaviour on the other.

TABLE 3

PEARSON'S R CORRELATIONS BETWEEN TEACHER RESPONSES (AS A PERCENTAGE OF TOTAL RESPONSES) AND PUPILS' ON-TASK BEHAVIOUR

Teacher Variable	Correlation with On-task Behaviour
Academic approval	0.3470**
Social approval	0.0983
Academic disapproval	-0.1744
Social disapproval	-0.3966**
Approval (acad. & social combined)	0.4032**
Disapproval (" " " ")	-0.4032**
Academic (appr. & dis. combined)	0.3250*
Social (" " " ")	-0.3250*

* P < 0.01; ** P < 0.001

There was a negative correlation between pupils' on-task behaviour and the absolute number of responses (of all types) that a teacher made. The correlation coefficient was -0.2663 ($P<0.01$), indicating an association between lower teacher response rates and higher on-task rates.

DISCUSSION

Many of the findings of this study confirm those of Wheldall *et al.* (1989) in their UK OPTIC study. The great variation in the behaviour of teachers and their classes are findings shared by both studies. The focus upon academic behaviour, and the emphasis on approval except in response to social behaviour, are as evident in the Hongkong research as in the Birmingham study. Strong positive correlations in this study between the use of approval (particularly academic) and on-task behaviour, and strong negative correlations between disapproval (particularly social) and on-task behaviour confirm the findings of the UK study.

The similarities between the UK and Hongkong findings are striking in view of the distinctive features of the Hongkong education system. One of these features, large classes, was apparent in this study, the figure for the Hongkong study being almost double that for the Birmingham research.

A number of differences between findings in the two studies should be noted. First, the level of on-task behaviour appears to be on average lower in Hongkong classes. This may be related to differences in the content of lessons and in teaching method, there being a great emphasis upon lecturing, rote-learning and preparation for in-school and public examinations in Hongkong schools (Morris, 1988). The presence in the class of pupils who are being made to repeat an entire year of schooling, or the use of English as a medium of instruction, may also have an effect. Differences in class size in England and Hongkong may be another factor, even though no relationship was found between the two variables within the Hongkong study itself.

Teachers in this study appeared to emphasise pupils' academic behaviours even more than their UK counterparts. Figures for the two studies were around 66 per cent and 60 per cent respectively, with corresponding figures of 34 per cent and 40 per cent respectively for social behaviour. The more pronounced focus on academic behaviour in Hongkong may be related to the priority given to academic success by individual teachers and the education system generally (Morris, 1983, 1988).

The emphasis upon approval (at around 63 per cent even more pronounced than that of 55 per cent for the UK study) is more surprising. Student teachers at Hongkong University sometimes suggest that the use of reinforcement for desirable behaviour is alien to traditional Confucian notions of the responsibility of teachers to instruct and pupils to study and obey without question and without requiring encouragement. The findings of this study suggest that, regardless of cultural traditions and some student teachers' notions of what may be appropriate, most teachers emphasise approval in ways consistent with the recommendations of modern Western behavioural psychology.

Further research using OPTIC is planned in Hongkong to investigate the variations individual teachers display in their patterns of response to different groups of pupils, and the relationships between patterns of teacher response and the expectations and attributions employed by teachers.

Correspondence and requests for reprints should be addressed to Sam Winter, Department of Education, University of Hongkong, Hongkong

REFERENCES

JOHNSON, R. K. (1983). Bilingual switching strategies: a study of the modes of teacher-talk in bilingual secondary school classrooms in Hongkong. *Lang. Learn. Communic.*, 2, 267-285.

HONGKONG EDUCATION DEPARTMENT (1986). *Annual Summary 1985-86.* H.K.E.D.

HONGKONG GOVERNMENT (1987). *Hongkong Nineteen Eighty Seven.* Hongkong Government Information Services.

MERRETT, F., AND WHELDALL, K. (1986). Observing Pupils and Teachers in Classrooms (OPTIC): a behavioural observation schedule for use in schools. *Educ. Psychol.*, 6, 57-70.

MORRIS, P. (1983). Teachers' perceptions of their pupils: a Hongkong case study. *Res. in Educ.*, 29, 81-86.

MORRIS, P. (1985). Teachers' perceptions of the barriers to the implementation of a pedagogic innovation: a South-East Asian case study. *Int. Rev. Educ.*, 31, 3-18.

MORRIS, P. (1988). Teachers' attitudes towards a curriculum innovation: an East Asian study. *Res. in Educ.*, 40, 75-87.

NAFPAKTITIS, M., MAYER, G. AND BUTTERWORTH, T. (1985). Natural rates of teacher approval and disapproval and their relation to student behaviour in intermediate school classrooms. *J. educ. Psychol.*, 77, 363-367.

WHELDALL, K., HOUGHTON, S AND MERRETT, F. (1989). Natural rates of teacher approval and disapproval in British secondary school classrooms. *Br. J. educ. Psychol*, 59, 38-48.

(*Manuscript received 7th December*, 1988)

Br. J. educ. Psychol., 60, 93-98, 1990

THE EFFECT OF ANIMATED DIAGRAMS ON THE UNDERSTANDING OF A MATHEMATICAL DEMONSTRATION IN 11- TO 14-YEAR-OLD PUPILS

BY S. V. THOMPSON AND R. J. RIDING
(Examinations and Assessment Research Unit, University of Birmingham)

SUMMARY. One hundred and eight children from 11 to 14 years were given instruction on shears and rotations, and how these transformations could be used to show the truth of Pythagoras's Theorem. The experimental group, which saw a computer-presented animated display of the transformations, scored more highly on a subsequent test than either of two control groups, the first of which saw no computer at all, but only a diagram sheet which was in the possession of all the children. The second control group saw a computer-presented display which showed only the same ten discrete stages of the transformations as the diagram sheet. It was concluded that animation per se can facilitate the understanding of a mathematical task.

INTRODUCTION

The microcomputer, with its capacity to produce high-resolution, animated graphics, presents an opportunity to overcome some of the limitations of conventional, paper-drawn illustrations of textual material. One of these is that such illustrations are static, although a sequence of diagrams may show the various discrete stages of, for example, a geometric transformation, leaving the pupil to fill in the gaps. An animation may be regarded as the limit of the possible number of stages shown, although in this case each previous state disappears as the next is drawn.

Conventional book illustrations have already been shown to be helpful, at least for the learning of the illustrated material (for a review, see Levie and Lentz, 1982), although the learning of non-illustrated material may also be facilitated by illustrations in factual texts (Rusted and Coltheart, 1979; Rusted and Hodgson, 1985).

Relevant, paper-drawn sequences of diagrams, then, might be expected to increase general understanding of a mathematical proposition, but an animated diagram should be superior as a learning aid as it maximises the amount of information given.

A number of investigators, however, have cast doubt on computer-based illustrations generally, and animation in particular, as an effective means of increasing the knowledge and understanding of a text. Reid and Beveridge (1985), presenting illustrated texts, found that learning scores were significantly higher in the traditional mode of presentation than in the computer mode. Phillips (1986) warned that animation can become a goal in itself and can force the viewer to structure information in a particular way. Rigny and Lutz (1976) found that interactive, animated graphics illustrating concepts in chemistry produced better knowledge and application, but the improvement in comprehension was not significant ($P<0.1$). This experiment, however, contrasted animated illustration with no illustration.

The aim of the present experiment, then, was to test the hypothesis that animation of diagrams would facilitate learning. As animation per se was being tested, an important aspect of the experiment was that, for educationally meaningful improvement to be claimed, the experimental group would have to do significantly better than two control groups: the first with a computer, but one which only presented a sequential series of static diagrams; the second with no computer at all, who went straight to the test after the lesson. Only if this double superiority occurred, could the value of animation be demonstrated.

METHOD

Sample

This comprised 108 children (54 boys and 54 girls) aged between 11 and 14. The sample was originally 201 children drawn from two suburban high schools and four suburban middle schools, and the results of 93 children were discarded at random within sexes to achieve equal numbers in the experimental and control groups, and in the three year groups (11+, 12+, and 13+). The large number of eliminated participants was due to extreme imbalances in the sexes in some mathematics groups in certain schools.

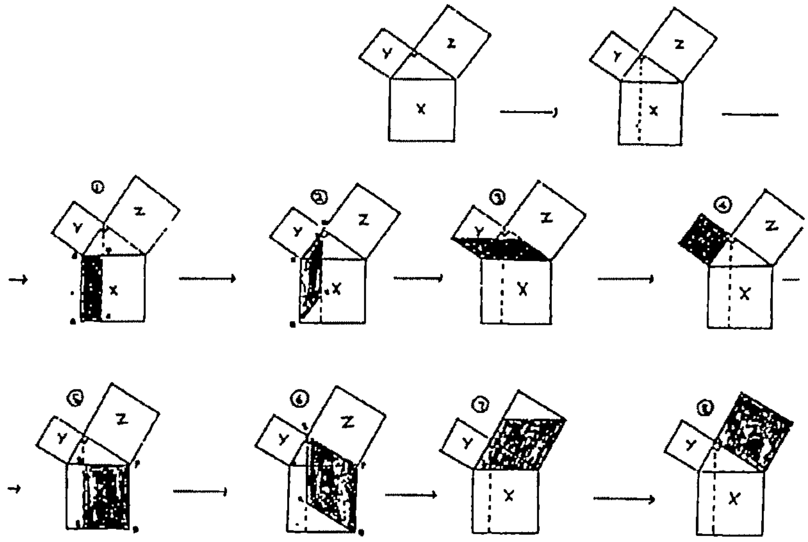
Materials

The materials comprised:

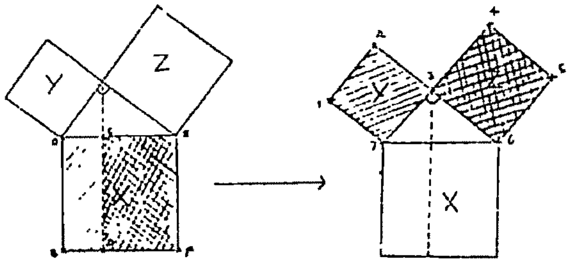
(1) Cardboard and other models as visual aids for a lesson on Pythagoras's Theorem. A pile of books with thin but visibly rectangular spines for demonstrating the shear transformation. Some of the models were mobile. These were two showing rotations, and a wooden rectangle which could be 'tilted', contrasting the reduction in vertical height and area with the constancy of the shear transformation in these respects.

FIGURE 1
THE DIAGRAM SHEETS

1.



2.



(2) Two sets of diagram sheets. The first, for general use by pupils throughout the experiment, comprised a sequence of 10 diagrams showing a demonstration of Pythagoras's Theorem using shears and rotations. This was adapted from a similar sequence in the Schools Mathematics Project (1972). The rectangle under transformation was shaded in brown with a green edge. The second set, for use with Test Two, showed just two diagrams with the same rectangles positioned before and after transformation. In each diagram, one rectangle was shaded red and the other yellow. The vertices of the two diagrams were labelled with letters and numbers respectively.

(3) Two sets of test papers. Test One consisted of 16 statements, each of which was unambiguously true or false. There were four statements each, concerning transformations, length, area and angle. For each category, two were true and two were false. The questions were arranged so that they had, as far as possible, the same mean position within the test without the categories ever rotating in the same order. The mean positions were 9.0, 8.0, 8.5 and 8.5 respectively. Further, the pattern of true and false statements was varied for each category. Test Two comprised six questions which could be answered from the second diagram sheet. Pupils were asked to which number each lettered point moved in the course of the transformations.

(4) Two versions of a specially written computer program, "Pythag" which demonstrated the truth of Pythagoras's Theorem using shears and rotations. Both versions showed an identical configuration to that at the top of diagram sheet one. Pupils then saw a dotted line gradually form. Version one changed discretely, flashing to the next situation after each completed transformation, and then executing a series of inverse transformations, thereby "undoing" the first series. Version two went through the same cycle, but the movement was comparatively continuous. The cycle took 2 minutes 40 seconds for each version.

PROCEDURE

Each session of the experiment was conducted using pre-existing mathematics groups, and because of the difficulty level of the task, no set lower than "three" was used. This was usually out of five sets. The pupils were therefore all considered able enough to gain sufficient grasp of the ideas being explained for any experimental effect to operate.

Balance for school was achieved across groups within each year group, but not between each year group. Such balance could not be attained as the age of transfer from middle to high school within the authority concerned was 13+, cutting across the age groups studied.

The experiment was carried out on each set separately in a double period normally reserved for mathematics. In each case, the experimenter was male, but he was introduced by the normal teacher for that lesson.

Firstly, all the pupils for that session were divided into three groups, control groups one and two, and an experimental group. Control group one was to see no computerised illustration, control group two was to see version one of "Pythag", which presented no more information than the illustrated diagram sheet, and the experimental group was to see the continuously moving version two (to control for Hawthorne effects, for spending more time on the lesson overall, but also for having a longer delay between the lesson and the test than control group one). All the pupils were then given a 15-minute lesson together by the experimenter, during which there was a brief discussion of rotations, shears and Pythagoras's Theorem, in very simple terms and with visual aids. No questions were asked or accepted. A set text was used.

After that, control group one went to a nearby room with their usual mathematics teacher and took the two tests, the introduction to these being given by the teacher, again according to a set text. Control group two and the experimental group were seated in the same configuration as each other around two computers, one at either end of the classroom, so that the two groups of pupils were facing away from each other. Each group saw one complete cycle of their version and were then told that the questions would be about the lines, angles and areas involved, as well as about the actual transformations themselves. The same hints were included in the test instructions so that control group one received them too.

The two groups then saw two further cycles of their program version. They then went back to their desks and took tests one and two for 15 and 10 minutes respectively. No talking was allowed throughout the whole period of the experiment. No questions were asked or answered except to ascertain whether pupils had understood the test instructions. If any indicated that they had not, the same instructions as before were read through again with the pupils. Pupils were not allowed to use rulers or other measuring devices during the tests.

FIGURE 2
THE TEST PAPERS

TEST 1

After each of the following statements, write 'T' if the statement is true and 'F' if the statement is false. Use the boxes provided.

EXAMPLE: In each of the transformations between 1 and 4, the corner 'B' of the brown region never moves F

If you wish to change an answer, put a cross over the whole box and write the new letter outside. ☒ F

1. The transformation in the brown, shaded area between 1 and 2 is the same type of transformation as that which takes place between 3 and 4. ☐
2. In the course of the first three transformations (between diagrams 1 and 4), all four sides of the brown, shaded area change their lengths at some stage. ☐
3. The areas of the brown, shaded regions in 1 and 2 are NOT equal. ☐
4. The interior angle of the corner of the brown shape at 'A' in 2 gradually increases as the region transforms to the brown shape in 3. ☐
5. In 5, the line RQ is shorter than any side of the square Z. ☐
6. The brown, shaded, rectangular part of the square X in 5 is smaller in area than square Z. ☐
7. In 6, the shaded corner at Q MUST fit into the corner of the triangle marked with a dot. ☐
8. The transformation in the brown, shaded area between 6 and 7 is a 'shear'. ☐
9. In each of the diagrams 5, 6, 7 and 8, the brown, shaded region has the same area. ☐
10. In the course of the transformation in the brown, shaded region between 6 and 7, all the angles of that region remain unchanged. ☐
11. In each of the diagrams 5, 6, 7 and 8, the green line is the same length. ☐
12. The eight diagrams show the rectangle ABCD in 1 gradually being transformed until it fills the square Z in 8. ☐
13. The diagrams 1 to 8 show that the interior angles of the brown, shaded region never really change although they appear to do so. ☐
14. Although individual sides of the brown region may change in length between 5 and 8, the overall perimeter of that region remains the same through each of the three transformations. ☐
15. Six transformations are shown, two rotations and four shears. ☐
16. The area of square Y, added to the area of square Z, equals the area of square X. ☐

TEST 2

The diagrams on the separate sheet of paper show two shaded rectangles which, together, make up square X (Figure 1).

These are then 'transformed', using shears and rotations, until they fill squares Y and Z (Figure 2).

In figure 1, six points, A, B, C, D, E and F have been marked with a cross and labelled with their letters.

As the rectangles change and move, so the points on them will change and move, and eventually come to rest at one of the numbered points shown in Figure 2.

By completing the sentences below, say at which number each lettered point will end up.

1. A will end up at .....

2. B will end up at .....

3. C will end up at ..... and .....

4. D will end up at ..... and .....

5. E will end up at ...

6. F will end up at .....

(Note that C and D lie on both rectangles and will therefore move to two different points.)

RESULTS

The means for each treatment group, out of a possible score of 20, (SDs in brackets) are:

GROUP 1 (Control — no computer)	10.58 (3.76)
GROUP 2 (Control — with computer)	11.04 (3.19)
GROUP 3 (Experimental group)	12.33 (3.51)

Direct, a priori comparisons were made between the experimental group and the two control groups (Winer, 1962, p. 208). The 0.05 rejection level was maintained as the experimental group had to score more highly than both control groups for the usefulness of animation to be established.

Group three was superior on a one-tailed t-test both to group one ($t = 2.09$, df 70, $P<0.05$) and to group two ($t = 1.75$, df 70, $P<0.05$).

DISCUSSION

The improvements due to animation were significant but small in educational terms. This may, however, have been due to the constrained nature of the experiment. In an ordinary classroom setting, the teacher may be able to increase the effectiveness of animated diagrams by discussing the transformations as they occur, whereas in the experimental situation the pupils had to sit and watch them in silence with only a prior warning from the teacher to direct their attention to salient points of the display. Also, the stage at which the animation is shown in the teacher's exposition could be important, and in a normal lesson the teacher would be free to vary this as s/he felt appropriate. Some of the fears expressed about animation, for example, that "one problem with animation is that viewers are largely deprived of the opportunity to structure the information for themselves" (Phillips, 1986), appear not to have been well founded in this sort of situation.

The reason for the improvement is impossible to state for certain from the results of the experiment. The cause may have been that more information was provided — the idea underlying the hypothesis — or it may have been motivational in that version two, the fully animated version, may have been more interesting to watch than version one, the discretely changing version, which kept the pupils waiting as it flashed from stage to stage of the transformations. Hawthorne effects are an unlikely explanation, as control group two and the experimental group were unaware of any difference between them and received ostensibly identical treatment. Whatever the reason, the controls in the experimental design imply that it was the animation itself which brought about the improvement and not some other aspect of the computer-presented package.

Among a number of questions that could be raised here, some repeat those raised by investigators with respect to conventional text illustrations, concerning optimum amount of colour and detail, generalisation of benefit to non-illustrated matter, differences with immediate and delayed recall, and sex and other individual differences. New ones concern speed and smoothness of motion, automatic running versus pupil manipulation, the facility to change the animation through the manipulation of variables, possible distortions of understanding through distracting features of the animation, the separation of the animation from teacher exposition and the constitution of suitable subject matter.

It is hoped that the findings of this experiment will give some additional support to those writing animated computer programs, the efficacy of which has hitherto only been assumed.

Correspondence and requests for reprints should be addressed to Dr S. V. Thompson, Examinations and Assessment Research Unit, School of Education, University of Birmingham, PO Box 363, Birmingham B15 2TT.

REFERENCES

Levie, W. H., and Lentz, R. (1982) Effects of text illustrations: a review of research. *Educ. Commun. Technol.*, 30, 195-232.

Phillips, R. J (1986). Computer graphics as a memory aid and a thinking aid. *J Comput assist Learn.*, 2, (1).

Reid, D. J., and Beveridge, M. (1985) Mode of presentation: effects on children's learning. Reported in Reid, D , and Shields, C. (1985), Microcomputers and biology software — identifying a problem. *Res. in Educ.*, 34, 1-8.

Rigny, J , and Lutz, K. (1976). Effect of graphic analogies of concepts in chemistry on learning and attitude. *J. educ. Psychol.*, 68, 305-311.

RUSTED, J. M., AND COLTHEART, M. (1979). Facilitation of children's prose recall by the presence of pictures. *Memory and Cognition*, 7, 354-359.

RUSTED, M. M., AND HODGSON, S. (1985). Evaluating the picture facilitation effect in children's recall of written texts. *Br. J. educ. Psychol.*, 55, 288-294.

SCHOOLS MATHEMATICS PROJECT (1972). *Teachers' Guide for Book G*. Cambridge: Cambridge University Press, p. T174.

WINER, B. J. (1962). *Statistical Principles in Experimental Design*. New York: McGraw-Hill.

(*Manuscript received 26th January*, 1989)

Br. J. of educ. Psychol., 60, 99-105, 1990

THE RELATIVE EFFECTIVENESS OF FOUR METHODS OF TEACHING PRACTICAL SCIENTIFIC SKILLS TO SECONDARY SCHOOL STUDENTS

BY G. M. SEDDON AND J. F. BARRY
(University of East Anglia, Norwich)

SUMMARY. The investigation compares the effectiveness of four methods of teaching students how to perform two practical scientific skills. The methods comprise commonly used combinations of four different modes of communication. Method W uses written instructions only. Method WG has these written instructions supplemented with graphics. Method WGS adds spoken instructions to the combination used in Method WG, and Method WGSD also includes a video-recorded demonstration of the practical task. After teaching random groups (N = 20) of students aged 14-15 years with one of the methods, their effectiveness was assessed in terms of the students' performance on test exercises which required the students to carry out these tasks. The results showed that the methods involving the use of spoken instructions (i.e. WGS and WGSD) were significantly worse than those that did not use spoken instructions. The use of graphics also had no significant beneficial effect, and the demonstration had a significant positive effect only in one task.

INTRODUCTION

One of the major innovations in all new and forthcoming science curricula is the specific inclusion of the use of measuring instruments and apparatus as tasks to be taught and tested. The current state of students' performance in these skills in the United Kingdom has been monitored by the APU surveys (e.g., Driver *et al.*, 1982) and the TAPS project (e.g. Bryce *et al.*, 1984). In each case, the various reports paint a surprisingly disappointing picture, even with students of age 15.

> "... knowledge of the basic practical skills is not as widespread as might have been anticipated."
> (Driver *et al.*, 1982, p. 64)
>
> "Even in the case of measurement skills — which are close to the hearts of many science teachers — performance is poor."
> (Bryce *et al.*, 1984, p. 65)

There is therefore an obvious need for detailed evaluation of the different methods for teaching these skills.

At the present time the methods most used for teaching practical skills involve four different modes of communication, i.e., written instructions, graphics, spoken instructions, and demonstrations. Moreover, they may be used alone or in various combinations. Unfortunately neither psychological theory (e.g., Hartman, 1961; Van Mondfrans and Travers, 1964; Travers, 1967; Paivio, 1971; Salomen, 1979; Kosslyn, 1980) nor empirical research provide any clear guidance as to the optimum way of combining them.

Different theories lead to different predictions. For example, the dual coding theory of Paivio (1971) regards information represented by both pictorial and verbal codes as being processed by functionally independent, though interconnected, cognitive systems. It contends that information represented by both image and verbal coding, for example, as being more powerful than that represented by just one code. In contrast, other theories (e.g., Travers, 1967) argue that the transmission of redundant information through more than one channel can create dysfunctional "jamming". Hence, the learning via two or more channels should not be superior to that from a single channel.

As indicated in the review by Levie and Lentz (1982), and elsewhere (e.g. Levin, 1973; Rowher and Harris, 1975; Nasser and McEwen, 1976; Maher and Sullivan, 1982; Nugent, 1982; Baggett, 1984), the empirical research has many examples of investigations to compare these combinations of modes of communication in teaching cognitive skills. However, there are conflicting conclusions, and no general pattern of results has yet emerged. Unfortunately, there has been very little research into their relative merits in teaching practical and manipulative skills. In one investigation with university students, Kempa and Palmer (1974) studied the effectiveness of these different modes of communication as three combinations which varied in terms of the amount of concrete and visual help provided. They concluded that demonstrations were more effective than written instructions, but that there was no further advantage to be gained by adding written instructions to the demonstrations.

The present investigation sets out to examine the effects of varying degrees of concrete and visual help with secondary school students. As well as studying two methods similar to those used by Kempa and Palmer, it also includes two other combinations of the four modes of communication to produce teaching methods which can be regarded as intermediate in terms of the amount of concrete and visual help. In Method W, the students are given written instructions alone. In Method WG, these same written instructions are supplemented with graphics to illustrate the performance of various operations. Method WGS employs the written instructions and graphics of Method WG, but while the students read the instructions themselves, they also hear them spoken. Method WGSD subsumes the combination of Method WGS, but the tasks are also demonstrated visually to the students at appropriate moments during the spoken commentary.

METHOD

The basic plan was to teach the same tasks using each of the four methods with different random samples of students taken from the same population. The effectiveness of the different methods was to be compared in terms of the performance of the different groups of students in carrying out these tasks in test exercises immediately afterwards.

Students

The students were all those taking fourth-year science courses in a suburban comprehensive school in Norwich. The 80 students, all of European extraction, were aged 14-15, and the two sexes were almost equally represented. They covered the whole range of ability, the average being close to that for the country as a whole.

Practical tasks

The two tasks concerned the use of measuring instruments and apparatus from the list given by Driver *et al.* (1982, pp. 203-205). One was the measuring cylinder: the students were to be taught how to measure a volume of liquid in the range 10-100 cm^3 correct to the nearest 1 cm^3 graduation and pour it all into a beaker. The other item was the Bunsen burner in its particular application concerned with the heating of solids in ignition tubes. The main points of technique to be taught were using the correct colour and area of the flame, as well as holding the tube at the correct angle in the flame.

Teaching materials

For both tasks the teaching materials were developed and subjected to pilot trials in a different school from that taking part in the main experiment.

Written instructions. Both the tasks were divided into sub-tasks. The instructions for executing these sub-tasks were then printed in order as short sentences on worksheets using language of a level which the teachers in the school agreed was appropriate. Examples of the instructions given for using the Bunsen burner are given in Figure 1.

Graphics. The graphics were printed at appropriate points in the right-hand margins of the worksheets and depicted the major features of good technique for each sub-task. As shown in Figure 1, all the pieces of apparatus were drawn in the conventional form used in standard text-books for this level.

Spoken instruction. The spoken instructions were given by means of an audio-cassette made in a professional recording studio. The words used were identical to those used in the written instructions.

Demonstrations. These were presented by means of a VHS video-recording made by professional staff in the television studios at the University of East Anglia. It depicted one of the present authors demonstrating each of the sub-tasks in sequence and speaking the same words used in making the written and spoken instructions. The programme was made in colour, and in actual use the cassette was played on standard school equipment.

Test exercises

With both tasks, the test exercises were presented to the students as a larger investigation with a wider educational purpose. However, the use of the right technique determined the accuracy of the final results or observations. Performance was therefore measured in terms of the accuracy of these results or observations.

Measuring cylinder. These test exercises were included as a set of measurements to investigate the effects of changing the concentration of reactants on the rate of precipitation of sulphur from sodium thiosulphate by dilute acid. Thus students were required to measure out specified volumes of dilute acid and sodium thiosulphate solution using separate measuring cylinders for each, and mix the two

FIGURE 1
Examples of the Written Instructions and Graphics in the Teaching Materials Developed for Using the Measuring Cylinder.

PROCEDURE

Keep the measuring cylinder containing 50 cm^3 of solution Z on a flat surface and fill carefully. The final level should be checked by having your eye level with the mark.

Place the conical flask containing 50 cm^3 of solution Z on the piece of white paper which has a cross marked on it.

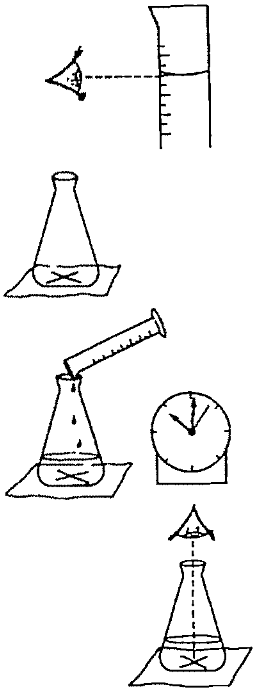
Carefully pour 5 cm^3 of the acid into a small measuring cylinder.

Quickly pour the 5 cm^3 of the acid into the conical flask containing the 50 cm^3 of solution Z.

Immediately start the stop clock. At the same time swirl the flask twice and place on the piece of white paper marked with a cross.

Look directly down through the open mouth of the conical flask so you can see the cross. Keep watching the cross until you can no longer see it.

Immediately note the time taken from the start until you can no longer see the cross.



solutions in a beaker. The speed of the precipitation was to be measured in terms of the time required for the disappearance of a cross drawn on a piece of paper placed underneath the beaker and viewed from directly above the liquid. For this purpose, the students were also provided with a standard laboratory stop clock which was to be started immediately the solutions were mixed.

This whole operation was carried out six times using a different concentration of acid in each case. In changing the concentration of the acid, the instructions required the students to dilute a specified volume of a stock solution with different volumes of distilled water. The volumes to be measured varied from 5-50 cm^3, and, with the concentrations used for the different solutions, the intervals of time ranged from 10-90 seconds.

A measure of the overall accuracy of any single time for disappearance was derived as its deviation from the corresponding time obtained by the present authors having previously carried out the whole set of measurements under the same conditions. For each concentration these deviations were then reduced to a five-point scale. For example, a time which deviated from the authors' by less than 10 seconds was given a score of 5; a deviation of between 10 and 20 seconds received a score of 4, etc., etc. Over all the different concentrations a combined measure of the accuracy was obtained as the mean of these deviations. Thus the greater is the numerical value of this measure, the better is the performance.

Bunsen burner. These test exercises were based on a set of experiments in which the students were to observe the effects of heat on the appearance of a solid. They were also to test for any oxygen and carbon dioxide given off in the process.

The accuracy of the students' results was determined from their written responses to the ten instructions listed in Table 1. The observations were then assessed on a three-point scale (i.e., 2, 1, 0) according to whether they were totally correct, partially correct or wrong, respectively. A combined measure was derived by adding the scores for all the responses. This test exercise was performed on five different solids, i.e., ammonium chloride, copper carbonate, lead nitrate, sodium hydrogen carbonate and hydrated copper sulphate. The overall assessment for all the substances was obtained by adding the measures for each substance.

TABLE 1

THE INSTRUCTIONS USED TO ELICIT THE OBSERVATIONS IN THE TEST EXERCISE ON USING THE BUNSEN BURNER

1. Describe the colour and general appearance of the substance before heating.
2. Describe the changes, if any, in appearance of the substance during heating. Mention any colour changes. If there was no change, write 'No Change'.
3. Describe the sound made, if any, by the substance during heating. If there was no noticeable sound, write 'No Sound'.
4. Describe the appearance of any solid substance forming towards the open end of the test tube during heating If there was no solid formed write 'None formed'.
5. Describe what happens when a glowing splint of wood is placed in the test tube during heating.
6. Describe what happens when some gas from the heated sample is passed into lime water and shaken.
7. Describe the appearance of any droplets of liquid which may have formed in the tube during heating.
8. Describe the colour of the liquid formed when the substance melted. If it did not melt, write 'The substance did not melt'.
9. Describe the appearance of any gas which you saw come off as the substance was being heated. Mention its colour. If no gas could be seen, write 'No gas seen'.
10. Describe the final appearance of the substance when the heating is stopped, and it is then allowed to cool down. Mention the colour.

Administration

The two tasks were taught to the same set of students during their normal science lessons in two successive weeks. The randomisation of the students to the four treatments was done separately and evenly for the two sexes and independently for the two tasks. In the main, students therefore followed a different treatment for each task.

Each of the methods was administered simultaneously in different laboratories, with the students working in randomly chosen pairs with each partner of the same sex. Each pair recorded the results or observations on specially prepared sheets, and submitted the sheet for assessment as a joint effort.

All the apparatus was already set out in designated work places before the students entered the laboratory. The lessons began with the minimum of routine spoken instructions from the supervisors, who were all experienced science teachers.

RESULTS

Table 2 summarises the measures of performance over all the pairs of students for both tasks. The main features are displayed to better effect in Figure 2. The results were analysed by a 4 x 2 analysis of variance, in which the factors were teaching method and sex, respectively, For both tasks the only significant differences were those relating to the main effect of teaching method ($P<0.05$).

These significant effects were examined in more detail by a series of planned orthogonal comparisons. For each task the combined mean of the two groups that did not have the benefit of any sound at all (i.e., those using Methods W and WG) was compared with the combined mean of the two groups that were taught using one form of sound (i.e., those using Methods WGS and WGSD). Further comparisons were made between the means from the Methods W and WG, and between the means from methods WGS and WGSD.

For both tasks there was a significant advantage in favour of those groups that had been taught without sound over those that had had sound ($P<0.05$). Only for the use of the Bunsen burner was there a significant difference between the means of the groups from Methods WGS and WGSD ($P<0.05$). The advantage was in favour of Method WGSD.

TABLE 2

Means (X) and Standard Deviations (SD) of the Different Treatment Groups on the Test Exercises

Test exercise		Teaching method*			
		W	WG	WGS	WGSD
Measuring cylinder	X	14.6	14.0	6.6	10.9
	SD	5.1	6.2	5.8	6.9
	N**	9	10	10	10
Bunsen burner	X	46.0	43.9	37.2	43.4
	SD	3.7	4.3	9.8	4.9
	N**	10	10	10	9

* W = Written instructions
G = Graphics
S = Spoken instructions
D = Demonstration

** Number of pairs of students

DISCUSSION

As shown in Figure 2, the means for the different groups range from those which are probably as bad as the performances noted by the APU and TAPS reports, to those which are as good as any experienced teacher might realistically expect from students covering the whole range of ability. In addition, the statistically significant differences confirm that the methods vary widely in effectiveness.

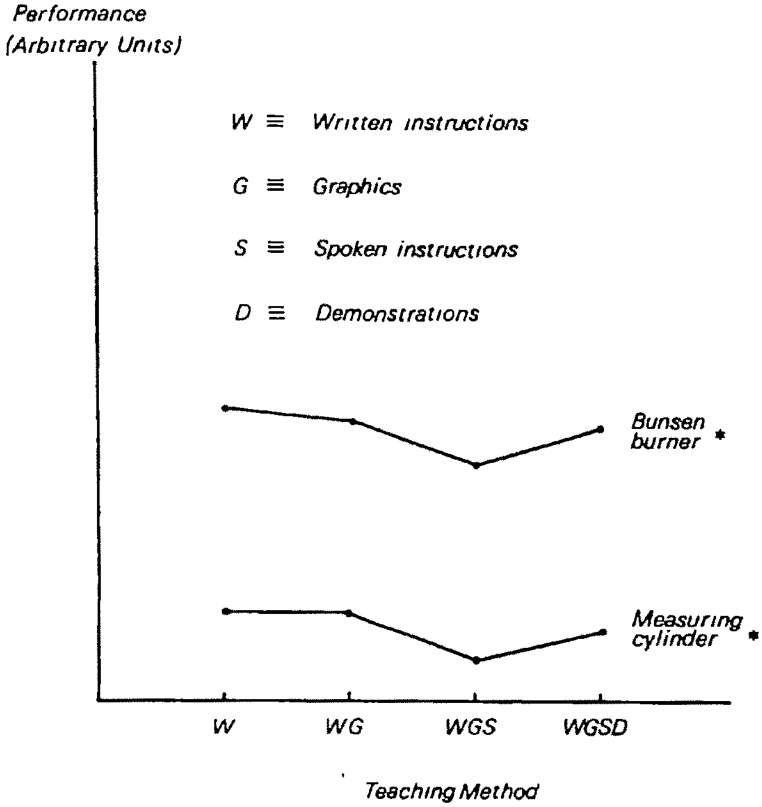
The direction of the significant differences quite clearly indicates that there has been a high degree of dysfunctional "jamming". In fact, those students who had methods involving two and three channels of communication did worse than those having methods with fewer channels of communication. Also, there were no significant benefits in adding either form of visual enhancement, i.e., the graphics or the demonstration. It is possible that, while listening to the spoken instructions, the students paid insufficient attention to the written instructions, and perhaps relied on their memory of what had been said, while performing the tasks. In contrast, the students who did not have the spoken instructions gained from the intellectual demands of assimilating the written word and became sufficiently familiar with the text to refer to it while carrying out the tasks. Whatever the correct explanation, the obtained result must be worthy of reflection for the very many teachers who often read aloud instructions which are also available in written form for their students. This well-meant additional help may actually be counter-productive.

The failure of the graphics to improve learning is less easily explained. While it is an adage to say that a picture is worth a thousand words, the present result implies that a picture adds nothing to the effectiveness of the words used in the present investigation. However, it is worth noting that, unlike the effects of the spoken word, the graphics did not actually do any harm. They just did not have any beneficial effects. Hence the expense in providing this kind of visual material is obviously wasted.

Any expectation that the method involving the demonstration would have been the most effective has certainly not been realised. Only in using the Bunsen burner is there any positive effect of using the demonstration on top of the combination WGS, but it does not make Method WGSD any better than Method W. This result is even more surprising, bearing in mind the finding of Kempa and Palmer (1974) that university students definitely gained from a demonstration given on top of written instructions. However, it must also be acknowledged that the design of the present experiment confounds any positive effects which the demonstration may have had with the clear negative effects observed for the spoken instructions. It is therefore possible that, if the demonstration had been presented silently as a supplement to the written instructions, the result could have been significantly better than that observed with Methods W and WG. Only further investigations can answer this question.

FIGURE 2

GRAPHICAL REPRESENTATION OF THE MEAN SCORES OF THE DIFFERENT TREATMENT GROUPS IN TEACHING EACH TASK



* The scales of measurement for performance are different for each test exercise

It is important to note that virtually the same result has been obtained with the two different tasks, and, although the students were largely the same in both cases, their random allocation to the different methods was completely different on each occasion. The two tasks used measures of performance which are also very different. In one case, they were quantitative and assessed objectively. In the other, they were qualitative and relied on the subjective impressions of human judges. Therefore, the rather surprising results cannot be an artefact of the choice of one particular task or method of assessment.

Finally, the investigation has been conducted with samples of a size comparable to that of normal classes in practical lessons in the United Kingdom. Thus the observed effects have been obtained in a situation which corresponds to conventional practice. In this sense, both the significant and non-significant effects observed can be regarded as being educationally significant.

Correspondence and requests for reprints should be addressed to Dr G. M. Seddon, School of Chemical Sciences, University of East Anglia, Norwich, Norfolk NR4 7TJ.

REFERENCES

BAGGETT, P. (1984). Role of temporal overlap of visual and auditory material in forming dual media association. *J. educ. Psychol.*, 76, 408-417.

BRYCE, T. G. K., MCCALL, J., MACGREGOR, J., ROBERTSON, L. J., AND WESTON, R. A. J. (1984). *Techniques for the Assessment of Practical Skills in Foundation Science: Report of the Project*. Glasgow: Jordanhill College of Education.

DRIVER, R., GOTT, R., JOHNSON, S., WORSLEY, C., AND WYLIE, F. (1982). *Science in Schools. Age 15; Report No. 1*. London: HMSO.

HARTMAN, F. R. (1961). Single- and multiple-channel communication: a review of research and a proposed model. *AV Commun. Rev.*., 9, 235-263.

KOSSLYN, S. (1980). *Image and Mind*. Cambridge, Mass.: Harvard University Press.

KEMPA, R. F., AND PALMER, C. R. (1974). The effectiveness of video-taped and recorded demonstrations in the learning of manipulative skill in chemistry. *Br. J. educ. Technol.*, 5, 62-71.

LEVIN, J. R. (1973). Inducing comprehension in poor readers: a test of a recent model. *J. educ. Psychol.*, 65, 19-24.

LEVIE, W. H., AND LENTZ, R. (1982). Effects of test illustrations: a review of research. *Educ. Commun. Technol. J.*, 30, 195-232.

MAHER, J. H. JR., AND SULLIVAN, H. (1982). Effects of mental imagery and oral and print stimuli on prose learning of intermediate grade children. *Educ. Commun. Technol. J.*, 30, 175-183.

NASSER, D. L., AND MCEWEN, W. J. (1976). The impact of alternative media channels: recall and involvement with messages. *AV. Commun Rev.*., 24, 263-272.

NUGENT, G. C. (1982). Pictures, audio and print: symbolic representation and effect of learning. *Educ. Commun. Technol. J.*, 30, 163-174.

PAIVIO, A. (1971). *Imagery and Verbal Process*. New York: Holt Rinehart and Co.

ROWHER, W. D., AND HARRIS, W. J. (1975). Media effects on prose learning in two populations of children. *J educ. Psychol*, 67, 651-657.

SALOMEN, G. (1979). *Interaction of Media, Cognition and Learning*. San Francisco: Jossey-Bass.

TRAVERS, R. M. W. (1967). *Research and Theory Related to Audio-visual Information Transmission*. (Rev. ed.). Kalamazoo, Michigan: Western Michigan University.

VAN MONDFRANS, A. P., AND TRAVERS, R. H. (1964). Learning of redundant information through two sensory modes. *Percept mot. Skills*, 19, 743-751.

(*Manuscript received 1st March*, 1989)

Br. J. educ. Psychol., 60, 106-115, 1990

CONFERENCE ABSTRACTS

Papers delivered at the Annual Conference of the Education Section of the British Psychological Society, held in the Durley Dean Hotel, Bournemouth, October 1989.

BLATCHFORD, P., Thomas Coram Research Unit, London

PLAYTIME IN JUNIOR SCHOOLS: THE CHILD'S VIEW

Playtime is a neglected but important part of the school day. In previous work at the Thomas Coram Research Unit (Blatchford, 1989) primary teachers' view on playtime, and children's behaviour then, were described. Teachers felt that aggressive and desultory behaviour were common, and that traditional playground games were in decline. But children are the experts on, and main witness of, playground life, and this paper reports on interviews with 11-year-old children in ILEA junior schools.

Playtime was very popular with children, particularly the long dinner break, and suggestions for improvements need to build on this positive view. But there were worries, for example, because of the cold, not knowing what to do, and because of disruptive behaviour. Just under one half of the children would have liked to stay in at playtime. It is argued that we take unacceptable playground behaviour too much for granted, and that children themselves would be responsive to constructive improvements. Children's own suggested improvements are considered.

The most common playground games were football, chasing games and other ball games. There were no reports of games with materials, like marbles. Singing and ring games, and games using playground markings, were rarely reported. It is argued that more could be done to encourage a greater variety of playground games. The paper ends with a look at the role of the most popular game of all — football.

CALDERHEAD, J., Department of Educational Research, University of Lancaster

STUDENT TEACHERS' CONCEPTIONS OF TEACHING, LEARNING AND THE CURRICULUM

Research on expert and novice teachers suggests that experts are distinguished by a rich, well-organised, episodic knowledge-base. The wealth of experience on which the experts can draw enables them to perceive classroom situations in a more differentiated way, to anticipate events more readily, and to adapt their behaviour accordingly. Student teachers have little practitioner experience on which to draw, yet at the same time they enter teacher education having spent thousands of hours in an "apprenticeship of observation". What, then, is the nature of the knowledge student teachers bring with them to teacher education? How does this knowledge develop? How, if at all, does it influence what students learn from their training experiences? What is the relationship between student teachers' entering conceptions of teaching and their developing classroom practice?

These questions were explored using a variety of methods to gain access to the conceptions of teaching, learning and the curriculum held by a sample of 12 student teachers. The students were interviewed about their experience at regular intervals throughout their first year. In order to tap their more tacit understandings of classrooms, videotapes of teaching were shown to the students who were invited to comment on what they observed; students were also asked to write accounts of lessons in the form of scripts as a means of making explicit their assumptions about teaching and children.

Analysis of the data suggests that student teachers have different expectations for, and approaches to, learning to teach. They also enter teacher education with several general images of teaching and of themselves as teachers. These images influence how they interpret and evaluate their experience in schools and what they extract from their teacher education courses. Different images appear to become modified and extended in different circumstances.

The research raises several questions concerning our theoretical appreciation of teachers' practical knowledge and professional learning, and illuminates some aspects of the nature of the images or mental models that are involved in teachers' professional thinking. The research also raises several questions concerning how teacher education courses might be most appropriately structured to challenge and extend students' understanding of classroom practice.

Creeser, A., Thomas Coram Research Unit, London.

BROS, THE BILL AND THEN TO BED: LOOKING AT 11-YEAR-OLD'S LEISURE ACTIVITIES

How do 11-year-olds spend their leisure time? There appears to be little quantitative work considering the range of leisure activities that this age group are involved in. Those studies which have looked at this area have tended to focus their attention on one or possibly two types of activities, most frequently television viewing, with the influence of television on children's behaviour receiving the most attention.

This paper gives an account of the ways in which a group of Inner London schoolchildren (N=71) in the final year at primary school reported spending their time. Each child gave an account of how they had spent the previous 24 hours when not at school. The discussion considers the amount of time that pre-adolescents devote to a range of educational activities including use of home computers for reading and maths; viewing television or video; listening to music and doing household chores. The extent to which the children's ethnic group and sex are associated with each of these leisure activities is explored.

Durndell, A., Siann, G., and Glissov, P., Department of Psychology, Glasgow College.

GENDER DIFFERENCES AND COMPUTING IN COURSE CHOICE AT ENTRY INTO HIGHER EDUCATION

Three hundred and eighty-seven students who had just entered higher education were asked why they had made their course choices, and what their views of the situation where low numbers of females were entering computing and technological areas. Whilst males in general were more instrumental in their course choices than females, computing students of both sexes appeared attracted by extrinsic awards perceived to go along with computing, and tended to deny that they were especially suited to study in their area. Potential female students of computing appeared to be put off by the prospect of harassment in predominantly male groups, and by problems linked to images of unfemininity. Whilst it was generally thought that females had the ability to study computers, school teachers were widely criticised for putting schoolgirls off technological subjects.

Entwistle, N. and Tait, H., Department of Education, University of Edinburgh.

APPROACHES TO LEARNING, EVALUATIONS OF TEACHING, AND PREFERENCES FOR CONTRASTING ACADEMIC ENVIRONMENTS

Previous research has demonstrated that the academic environments provided by departments in higher education have direct effects on students' approaches to studying. But other studies have indicated that these effects are mediated by the students' own perceptions of those environments. Here two studies are reported which explore the relationships between approaches to learning, or study orientations, and perceptions of the academic environment. Those perceptions are measured in two distinct ways, one which minimises the effects of differential perceptions, and one which highlights them. Factor analyses of the responses of three groups of students taking engineering and psychology are used to clarify the nature of the relationships between study orientations and perceptions of the academic environment. It is found, as in earlier studies, that there are relationships which associate deep approaches with perceptions of relevance, and surface approaches with a heavy workload. But here it is also shown that students with contrasting study orientations are likely to define effective teaching in ways which reflect those orientations. Implications both for the design of feedback questionnaires and for the improvement of teaching and learning in higher education are discussed.

FAWCETT, A., AND NICOLSON, R., Department of Psychology, University of Sheffield.

AUTOMATICITY — A NEW FRAMEWORK FOR DYSLEXIA RESEARCH?

The performance of a group of 23 13-year-old dyslexic children was compared with that of same-age controls on a battery of tests of motor balance. A dual task paradigm was used — subjects performed each test twice, once as a single task, and once as a dual task concurrently with a secondary task. Two alternative secondary tasks were used, the classic counting backwards task and an auditory choice reaction task. Both secondary tasks were calibrated for each subject to ensure that their performance on the secondary task alone fell between pre-specified performance criteria. In all single task conditions there was no difference between the performance of the two groups. By contrast, in 19 out of 20 tests performed under dual task conditions, the dyslexic group were significantly impaired, whereas the controls showed no impairment, thus resulting in significantly better performance by the control group than the dyslexic group. The sole exception was that the dyslexic children were not impaired on the easiest balance condition with the choice reaction task. Under the dual task condition the dyslexic children also performed worse than the controls on the secondary task. It is very hard to accommodate the findings within the traditional framework of a deficit specific to lexical skills. One plausible explanation of the results is that, unlike the controls, the dyslexic children need to invest significant conscious resources for monitoring balance, and thus their performance is adversely affected by any secondary task which serves to distract attention from the primary task. This need for "conscious compensation" suggests that for dyslexic children the skill of motor balance is poorly automised. It is possible, therefore, that many of the reading deficits of dyslexic children are merely *symptoms* of a more general learning deficit — the failure to fully automise skills.

FARQUHAR, C., Thomas Coram Research Unit, London.

SAFER TEACHING OF SAFER SEX? PRIMARY TEACHING IN THE AGE OF HIV/AIDS

In recent years, increasing attention has been paid, both by researchers and educationalists, to the understanding and attitudes of secondary school children towards HIV/Aids. However, little attention has been paid to the implications of HIV/AIDS for primary school children. This paper discusses some of the political, moral and legal questions raised when discussing HIV/AIDS with children of primary school age. It addresses four main questions. Firstly, is HIV/AIDS of relevance to primary aged children? Secondly, what do 8- and 10-year-old children understand and feel about HIV/AIDS? Thirdly, what role could primary teachers play in helping children to further their understanding? And fourthly, how prepared are teachers for this role? The paper draws on a one-year pilot study carried out at Thomas Coram Research Unit (London Institute of Education), and funded by the ESRC. The research was concerned with the feasibility of carrying out research into young children's AIDS-related knowledge. It was based predominantly in one suburban primary school, and involved interviews with LEA officials, school governors, teaching and non-teaching staff and parents, as well as data collection with children.

GALLOWAY, D., Department of Educational Research, University of Lancaster

WHOSE SPECIAL EDUCATIONAL NEEDS SHOULD WE BE DIAGNOSING?

This paper argues that children's special educational needs cannot be seen in isolation from the professional development of their teachers. Attention is drawn briefly to the limitations in some conventional approaches to the assessment of behavioural problems. These are discussed in the light of recent research on school and teacher effectiveness. The paper identifies factors within schools which contribute to pupils' special needs. It is argued that these factors are most usefully conceptualised as the legitimate needs of teachers, and that the special educational needs of children result more frequently from failures to meet the needs of their teachers than from cognitive, personality or familial problems. It is argued that the principles of diagnostic teaching from a logical extension to criterion referenced assessment and apply as much to learning as to behavioural problems. As important, they can accommodate a range of theoretical orientations.

GALLOWAY, D., Department of Educational Research, University of Lancaster.

DISRUPTIVE BEHAVIOUR IN SCHOOLS: THE RELEVANCE OF MOTIVATION STYLE

This paper reviews briefly the evidence that schools exert a substantial influence over their pupils' behaviour. It also considers the limitations inherent in the school effectiveness literature on process variables associated with effective/ineffective schooling. It is suggested that teachers and researchers have neglected the importance of motivational variables as influences on pupils' *and* teachers' behaviour. What is needed is a perspective that links policy and practice at school and classroom level with the perceptions and motivational styles of individual teachers and pupils. Within the framework of research on motivational style, it is argued that the concepts of learned helplessness and of self-worth concern have particular relevance in understanding disruptive behaviour in schools.

GLISSOV, P., Department of Psychology, Glasgow College.

THE GENDER GAP IN COMPUTER USE: PUPIL MOTIVATION AND ATTITUDES IN SECONDARY EDUCATION

A questionnaire was administered to 429 pupils in five Scottish secondary schools in order to survey their beliefs about, and attitudes to, computer use, both within and without the school setting. Two issues are reported on in this paper: first the effect of the sex composition on the school (whether single sex or co-ed) was considered. The three schools surveyed here were fee paying "private" schools chosen for their homgeneous pupil intake, thus controlling for social status. In line with previous research it was found that the co-ed situation promoted a polarisation of attitudes, i.e., there was a tendency for girls and boys in a mixed sex situation to answer more differently from each other in comparison with the girls and boys in the single sex group. Secondly, the effect of the socio-cultural background of the schools' pupils intake ("private", "intermediate" and "disadvantaged") was investigated. All schools in this analysis were co-ed. Attitudes to computers were generally positive. However, the pupils from the "disadvantaged" area showed a more enthusiastic and to some extent less critical attitude to computer use than pupils in the other schools.

Looking at the sample as a whole, boys showed a higher degree of sex stereotyped attitudes in connection with computer use, and they reported more knowledge and experience of computer programming and games. Also in line with existing literature, the girls adopted an "I can't, we can't" approach to computers, i.e, they believed that while women as a whole could cope well with computers, as individuals they felt anxious.

Interpretations of the documented gender gap will be presented and some questions for future research highlighted.

GLYNN, E., Education Department, University of Otago

DISCIPLINE IS FOR THE WHOLE SCHOOL: COMMENTS ON THE REPORT OF THE ENQUIRY CHAIRED BY LORD ELTON

Successful behaviour management and disciplinary practices in a school require a school-wide policy and school-wide commitment. This paper draws out four areas of practical concern arising from the recommendations made in the Elton report. These are: (1) changes in staff behaviour required within a school-wide approach; (2) input from a curriculum of personal and social education; (3) relationships with parents and parent involvement; (4) integration of children with disabilities and special needs.

Within an effective whole school approach to behaviour, management and discipline, it will become increasingly difficult for individual teachers to justify excluding from their classes children who have disabilities, handicaps, or who experience behavioural and learning difficulties.

HASTINGS, N., Department of Educational Studies and Management, Reading University

MAKING SENSE OF THE SENSE TEACHERS MAKE OF EFFECTING CHANGE IN CHILDREN'S BEHAVIOUR

The Elton Report and the research which it commissioned confirm that the most frequently encountered problems of classroom behaviour are, individually, relatively trivial. They are problems of talking, interfering with the work of others etc. There is now a substantial body of evidence which identifies elements of teachers' strategies and school organisation and as being related to the frequency of such behaviours. What we lack within educational psychology is any coherent attempt to develop a framework within which to make sense of these findings.

Teachers who have made positive use of this evidence rarely fail to make sense of the effects of their revised strategies on individual pupils and whole classes. They frequently locate the changes in their "relationships" with pupils. They see the change as entailing more than a change of behaviour.

This paper examines the relationship between a number of theoretical constructs which embody the idea of felt personal influence (viz. Locus of Control, Attribution Theory, Self Efficacy, Learned Helplessness and Personal Construct Theory), the notion of "good relationship" and behavioural change strategies now commonly employed in schools.

HORNBY, G., School of Education, University of Hull

EFFECT ON FATHERS PARENTING CHILDREN WITH DOWN'S SYNDROME

A considerable amount has been written about the effects on mothers of parenting children with handicapping conditions, but very little on fathers. A review of the literature on fathers reveals a fairly consistent set of negative assumptions about their response to the situation. These assumptions appear to be based on limited data of doubtful reliability. In order to remedy this situation a descriptive study was conducted with a representative sample of fathers of children with a specific handicapping condition.

Ninety-seven fathers of children with Down's Syndrome were interviewed and 88 of these completed and returned a booklet of questionnaires. The interview was minimally structured in order to obtain the fathers' perspectives of their situation (qualitative data). The booklet contained questionnaires designed to measure the fathers' stress level; adaptation; personality; marital satisfaction; social support; family relationships; and, demographic variables (quantitive data). Analysis of qualitative and quantitative data is underway and preliminary findings will be reported. Implications for professional practice and for future research with families of children with handicapping conditions will be discussed.

JONES, D., CASBY, W., KUGLER, B. AND WATKINS, B., Birkbeck College, Harperbury Hospital, London.

A LONGITUDINAL STUDY OF COGNITIVE DEVELOPMENT AND ACADEMIC ATTAINMENTS IN DOWN'S SYNDROME CHILDREN

Thirty-six Down's Syndrome children, half of whom attended mainstream primary or nursery schools and half school for children with moderate learning difficulties, were assessed over a two-year period. Baseline and six-monthly measurements were taken of their progress in expressive language, comprehension, numeracy, verbal fluency, drawing ability and reading. Mental age was estimated at baseline and at the end of the two years. The mainstream children made significantly greater progress in numeracy, comprehension and mental age and compared favourably on all other measures. The results are discussed in terms of their implications for integrated education.

/

KNIGHT, P., St Martin's College of Higher Education, Lancaster

PRESCRIPTIVE AND PRACTICAL VIEWS OF CHILDREN'S LEARNING IN PRIMARY SCHOOL HISTORY

Educational practice and educational prescriptions necessarily — albeit often unwittingly — embody beliefs about how children learn and develop. One set of problems arises when mismatches occur: for example, between these beliefs and the espoused aims of teaching; or between the theories underlying the prescriptions and those underlying the practices.

In this paper the logic of a recent, detailed curriculum prescription (the interim report of the History National Curriculum Working Party) is examined, and compared to the views of children's development and learning held by 28 primary school teachers identified as skilled in history teaching. Some implications are noted for school and staff development, and for the implementation of the National Curriculum.

KRUGER, C., SUMMERS, M., AND PALACIO, C., Oxford University, Westminster College

LEARNING CONCEPTUAL SCIENCE — HELPING THE PRIMARY SCHOOL TEACHER

The Primary School Teachers and Science (PSTS) Project is concerned with developing INSET materials to help primary teachers improve their understanding of the science concepts they will have to deal with when teaching the National Curriculum. Taking to heart Ausubel's well-known dictum "the most important single factor influencing learning is what the learner already knows — ascertain this and teach him accordingly", the first phase of the Project's work involved in-depth interviews with primary teachers to identify their *existing* conceptions in the areas of *forces and their effects, changes in materials* and *energy*. Building on the results of the interview, survey methods were used to establish the *prevalence* of views differing from the accepted scientific interpretations, in a wider sample of primary teachers. Having established the prevalence of these "misconceptions", the project team is in a position to decide the areas where help is most needed. The next stage of the work involves the design and trial of INSET materials that starts from teachers' existing conceptions and, within a constructive view of learning, will provide appropriate experiences to help them move towards accepted scientific interpretations of the concepts involved.

The presentation will describe some of the findings of the interview and prevalence phases of this work, and will go on to outline some of the planned approaches to INSET that we think will help primary school teachers in the task of teaching science at this level.

MOONEY, A., Thomas Coram Research Unit, London.

TEASING AND FIGHTING IN THE JUNIOR SCHOOLS — WHAT CHILDREN HAVE TO SAY

The Infant School Project at Thomas Coram Research Unit followed the progress of all September entrants to reception classes in 33 ILEA schools, until the end of their first year as juniors (*Young Children at School in the Inner City*, by B. Tizard *et al.*, 1988). When aged 7, a subsample were interviewed about school. This paper examines what children have to say about teasing and fighting at school during interviews for a follow up study of the sample at 11. It would seem that teasing is a common feature of school life in that almost all of the 175 who were interviewed said that children at their school were teased and about two-thirds said that they themselves had been teased during the fourth year. But how and why are these children teased and what do they do about it? Who get involved in fights and why? How do children generally feel about fighting? The paper addresses these questions, with reference to sex, ethnic group and age differences. Some implications are noted for school and staff development, and for the implementation of the National Curriculum.

PAIN, J., Educational Psychologist, Lancaster.

UNDERACHIEVEMENT AND UNSETTLED BEHAVIOUR

Through whole-based surveys of children's writing from dictation of "standard sentences" it has been possible to quickly identify the incidence and location of learning difficulties. In "reply" to the findings, strategies and resourcing have been developed in discussion with the schools and Local Education Authority. In parallel there has been some use of a behavioural questionnaire through which teachers have been invited to rate their pupils in terms of memory, attention and conduct (M.A.C. Questionnaire).

Substantial associations are repeatedly found between pupil underachievement (as assessed through writing from dictation) and teacher rating of unsettled behaviour. In conjunction with surveys of literacy in special schools, meeting the needs of pupils with emotional and behavioural difficulties, strong links are presented between "unmet" learning needs and behavioural difficulties.

The philosophy of the strategies is to assume that learning difficulties in basic subjects beget behavioural problems, rather than presume, as teachers often do, that behaviour difficulties are the "cause" of underachievement.

A series of whole school surveys are discussed to back up this strategic philosophy.

Plewis, I., Thomas Coram Research Unit, London

CONTINUITY AND CHANGE IN SCHOOL ATTAINMENTS IN THE PRIMARY YEARS

This paper reports on a follow-up, to the end of junior school, of the cohort of children who, in 1982, entered 33 multi-ethnic infant schools from nursery school. The results for the infant school period were given in the book by Tizard *et al.* (1988). Here we look at attainment at the end of junior school, progress during junior school, and progress throughout the primary years. Only 149 out of the starting sample of 343 children were still in their original schools by the end of juniors. In addition, pupils who had moved but stayed within the London area were tested to give a final sample of 166, 59 of whom were black (Afro-Caribbean) and 78 were white. Group reading and maths tests were used, the Suffolk reading test and the Young's "Y" maths test. The pattern of results found during the infant school period was repeated for junior school, with girls, particularly white girls, ahead in reading, and white pupils, particularly white boys, ahead in maths. However, the mean standard scores for the sample were low — 90 for reading and 86 for maths. The correlations for reading and maths over time were substantial. The paper concludes with a comparison of the results from this study and the recent ILEA study of junior schools (Mortimore *et al.*, 1988), and makes suggestions for future research in this area.

Pumphrey, P., University of Manchester.

INTEGRATING THE TESTING AND TEACHING OF READING IN THE PRIMARY SCHOOL: CHALLENGES AND RESPONSES

This paper has two aims. First, to identify challenges facing both teachers and educational psychologists in integrating the teaching and testing of reading. Second, to consider some techniques that show promise and merit development.

The Report by the Task Group on Assessment and Testing (DES, 1988), "The National Curriculum (English)" (DES and the WO, 1989) and Draft Circular/89 "Assessments and statements of special educational needs: procedures within the education, health and social services" (DES, 1988) set the context within which the above issues will be addressed.

Challenges considered will include points deriving from theoretical and practical issues concerning the nature of reading development, the uses and abuses of reading tests and assessment techniques, pedagogy and the initial and INSET training of professionals.

Responses identified include an overview of reading test and assessment techniques that integrate the teaching and testing of reading. Three of these will be described in more detail. Recent research results will be presented and their implications for furthering co-operation between the educational psychologist and the class teacher will be considered.

Ridgway, J., Department of Psychology, University of Lancaster

INFORMATION TECHNOLOGY IN PRIMARY SCHOOLS

What experiences should primary school pupils receive regarding information technology (IT), and what are the barriers to providing them?

This paper will begin with a review of the attainment targets relevant to IT defined in National Curriculum documents. Perhaps the most obvious barrier to innovation is the supply of hardware. Such an analysis ignores a more fundamental problem of teacher education in the educational uses of computers: problems of teacher education will provide the focus for this paper.

A major source of classroom based support for practising teachers is provided by Advisory Teachers (ATs) who are funded under the Educational Support Grant scheme. Lancashire LEA, Cumbria LEA and Lancaster University are collaborating on a project funded by the Training Agency which aims to support ATs in their work. Some of the problems they encounter (and their successes) will be described.

A second group of teachers to be considered is teachers undergoing initial training. One might consider these people to be heralds of the brave new world of computer assisted learning. A survey of student teachers' uses of IT during their teaching practice, and of their views concerning IT will be reported. The survey does not support the "heralds" view!

The paper will conclude by highlighting ways in which psychological research might help teachers to meet some of the challenges posed by the National Curriculum, regarding IT.

SIANN, G., MACLEOD, H. AND GLISSOV, P., Glasgow College and University of Edinburgh

THE EFFECTS OF COMPUTER USE ON GENDER DIFFERENCES IN ATTITUDES TO COMPUTERS

The attitudinal consequences of a focussed LOGO programming experience which took place over the course of one term within 114 primary school children attending four Edinburgh schools is discussed with particular reference to gender differences. Half the pupils used list processing and half the pupils used graphic processing. In addition half the pupils worked in the same sex dyads and half in mixed dyads. Substantial differences between boys and girls were shown in their pre-experience attitudes towards computer technology. In general the boys appeared more confident of their personal ability in the field of computing. They also showed more interest in computing and were more likely to believe that computing required high ability. General attitudes to computer use were similarly positive for both sexes. After the computer experience gender differences in attitudes to computing diminished. The exception to this general tendency was that girls' anxiety levels relative to boys increased. These results did not appear to be related to either the type of activity the pupils were exposed to or the composition of working dyads. Observation of the pupils at the console did indicate that in mixed dyads boys dominated. Interpretations of the gender differences are presented and important questions for future research and practice are raised.

SIMPSON, M., AND SAUNDERS, A., Northern College of Education, Aberdeen Campus.

PUPILS' EDUCATIONAL EXPERIENCES — COHERENT, CONTINUOUS AND PROGRESSIVE?

The Scottish Education Department has funded a longitudinal study of pupils' learning experiences in late primary and early secondary school, for 1989-1992. The research proposal emerged from a recommendation made in the final report of the 10-14 committee of the CCC, convened to consider pupils' learning experiences in the transition period of late primary and early secondary education. The committee members stated that:

"We take it as axiomatic that a young person's experience of education should be coherent, continuous and progressive" (10-14 Report)

The concern of our study is consideration of whether pupils' experiences are as recommended in relation to:

(1) the account taken of pupils' understanding, knowledge and skills in curricular planning;

(2) what might constitute a progression in a variety of curricular areas;

(3) what is continuous in relation to the content and skills which comprise a progression within a particular subject area;

(4) what constitutes a coherent experience as regards classroom pedagogy.

In order to monitor pupil experiences, we are following a small sample of pupils over a four-year period. We will build up a description of individual experiences using a case-study method. Our starting point will be to investigate the interactive processes between pupil, teacher and curriculum, considering how systems are regulated through the action of various individuals (Giddens, 1979).

Our data will be collected from a variety of sources, i.e., assessment of pupil understanding in a number of conceptual areas of the curriculum, existing documentation, participant observation and interviews. In addition we will make recourse to substantive, formal and meta theory, in order to analyse our data.

SMITH, L., Department of Educational Research, University of Lancaster.

PRIMARY TEACHERS' BELIEFS ABOUT CHILDREN'S DEVELOPMENT IN HISTORY TEACHING

A commonly used argument states that, since good practice exists, it should be codified and broadcast. This argument is central to an on-going ESRC study of the primary history classroom, where a sample (N = 28) of teachers, known to be doing good history work, is currently under study. The present paper examines four issues: (1) the extent to which teachers are sensitive to differences in children's age in relation to their abilities; (2) teachers' diagnostic competence; (3) teachers' beliefs about the place of children's errors in history learning; (4) the extent to which developmental principles are reflected in teachers' beliefs and practices (Smith, 1987). The discussion focuses on

— the extent of the correspondence between teachers' beliefs and their teaching

— the basis of teachers' judgments and actions

— issues, both theoretical and methodological, arising in the documentation of good practice.

REFERENCE

SMITH, L. (1987). Developmental theory in the classroom, *Instruct. Sci.*, 16, 307-317.

WATT, H. D., Department of Computing Science, University of Glasgow.

ATTITUDE CHANGE IN SCHOOLGIRLS TO COMPUTERS

Due to the national decline in the percentage of women students entering higher-education computing courses, several UK universities and colleges have run computing workshops for schoolgirls. These workshops aim to show the young women who come along that there is an important place for them in computing.

Although total student numbers are increasing, industry is increasingly unable to find sufficient candidates to fill the available posts in computing and information technology. Therefore, it is important that girls should be made aware of the opportunities which computing can offer. Running workshops for schoolgirls is one way of increasing awareness and counteracting the "boys' toys" image, which has become prevalent over the last few years.

In 1988 Glasgow University ran its first "Women into Computing" workshops for third-year schoolgirls from Strathclyde, Central, and Dumfries and Galloway regions. More than 2000 girls from 125 schools applied for the 100 available places. These schools encompassed both the public and private sectors, and both girls' and mixed schools.

Therefore in 1989, Glasgow University's Computing Science Department worked together with industry to provide five one-day "Women into Computing" workshops. On each of the five days 20 third-year schoolgirls attended. Although this year 1000 girls from 110 schools attended a workshop, there were still some schools left on the waiting list. Also, more than 1500 girls were disappointed.

The 1000 girls took part in computer workshop sessions, saw a video on IT, and heard women from industry talk about their experiences and the opportunities open to girls who study computing. Each girl was given a questionnaire to complete and return at the beginning of the day and another questionnaire during the final session. Some questions were included in both questionnaires. The initial analysis of the girls' responses shows that a significant attitude change took place during the course of the workshop.

The success of these "Women into Computing" workshops indicates that providing a "girls only" environment is one way of heightening girls' interest in the subject.

WHELDALL, K., AND MERRETT, F., Centre for Child Study, University of Birmingham.

POSITIVE TEACHING: EFFECTIVE CLASSROOM BEHAVIOUR MANAGEMENT WITH PARTICULAR REFERENCE TO THE ELTON REPORT

Problems of discipline and classroom control have been perennial concerns for teachers but academic educationists and teacher trainers have largely ignored these issues, focusing instead on such matters as curriculum design and content. Little or no specific training in managing classroom behaviour is provided for teachers at either pre- or in-service level, as the recently published Elton Report on "Discipline in Schools" makes clear. The Secretary of State for Education has accepted two of the main recommendations of this report: "that in future all initial teacher training courses should include separate, compulsory elements of practical training in how to manage pupil behaviour" and "that the management of pupil behaviour should become a national priority in the Local Education Authority Training Grants Scheme for in-service training of teachers".

Our research programme based at the Centre for Child Study has been addressing precisely these issues, including the development of specific, research based, training courses for teachers in "Positive Teaching: effective classroom behaviour management", for both primary and secondary teachers. These courses have been well received by the teaching community for over five years and rigorous experimental evaluations have clearly demonstrated their effectiveness in changing teachers' classroom behaviour and, as a result, improving the behaviour of their classes.

WHITTINGTON, J., MRC Applied Psychology Unit, Cambridge.

LARGE VERBAL-NON-VERBAL ABILITY DIFFERENCES AND UNDERACHIEVEMENT

In a large-scale study (over 10,000 11-year-olds) of verbal/non-verbal ability differences and underachievement it was found that children who had large V-NV ability differences were more likely to be underachieving in reading comprehension or mathematics or both, than children of the same overall ability with more similar V and NV components.

The rates for "severe" underachievement (two standard errors below expectation, based on regression of achievement scores on ability scores) were two to ten times greater among the large V-NV ability differences group of children, nearly 20 per cent of whom were severely underachieving in reading comprehension, as were over 5 per cent in mathematics.

WILKINSON, J. E., Department of Education, University of Glasgow.

PSYCHOLOGY AND INITIAL TEACHER TRAINING: A RESPONSE FROM TRAINING INSTITUTIONS

Over the past decade the position and status of psychology in teaching training courses has been gradually eroded to a point where considerable concern is now felt amongst psychologists that psychology is no longer playing the role it should be in the preparation of teachers. In order to reverse this trend, it is considered desirable that the Society use its resources to bring together psychologists with important contributions to make to teacher training and to use the common ground to promote a defined curriculum process for institutions involved in teacher training.

In 1988 the Society published a Statement on Psychology and Initial Teacher Training (a summary appeared in the February 1989 issue of *The Psychologist*). The paper was prepared by an ad hoc working party of the Education Section and was approved by the SAB.

The statement was sent to all establishments (universities, polytechnics and colleges) involved in the training of teachers, together with a letter from the President of the Society requesting a response to the paper. Responses have now been analysed and the results of this analysis will be presented.

BOOK REVIEWS

AINSCOW, M., AND MUNCEY, J. (1989). *Meeting Individual Needs.* London: David Fulton, pp. 149, £10.95, ISBN 1-85346-033-8.

This book comes at a time when schools are undergoing great changes following recent legislation, and offers a way of viewing children and their education that is refreshing, and which redresses some of the unfortunate negative aspects of the 1981 Education Act. "The aim of this book is to help teachers in primary schools to review and develop their thinking and practice in meeting the individual needs of their pupils."

The authors call upon their considerable experience in working on SNAP (Special Needs Action Programme), a major in-service initiative based in Coventry. This initiative has influenced the response of many other local authorities to special needs provision.

Each chapter begins with a short chapter guide which helps focus the reader on the overall nature of the content. This I feel is particularly useful where the authors are dealing with content and philosophy simultaneously. Chapter One is an overview of Warnock and there is an interesting post-Warnock analysis which sets the scene for the rest of the book: "... we have argued that a post-Warnock era demands a change of emphasis away from a concern with a separate group of children, said to have special needs, towards an orientation that attempts to make schools more effective in responding to pupils as individuals."

Chapter 2 gives a clear and concise account of the SNAP initiative including evaluation data, the implications of school change, and an overview of the characteristics of those schools dealing effectively with special needs as a whole school approach. This leads directly into Chapter 3 which expands on how effective schools come to be effective and describes the necessary changes in individual teacher practice and school response. This chapter, "Developing Practice", was the most interesting for me because the authors have not glossed over the difficulties that change presents in a complex situation such as a school. The section on the nature of change is positive and encouraging, taking into account the human factors involved. I particularly welcomed the emphasis that "change is really about learning" and "change is a process not an event". There is a brief section on stress and I wanted this to be longer but, given the breadth of the book, it would have led to an imbalance. It is heartening however to see it included.

The chapter on "Assessment, Recording and Evaluation" does not address the National Curriculum directly but the framework offered is not incompatible with the general ethos of information-gathering implied in the National Curriculum. My only fear is that teachers will adopt less useful methods of assessment than those the authors offer. They offer a brief section on audio-visual recording which I feel may offer an exciting and informative way of recording children's hearing behaviour particularly for the practical aspects of the National Curriculum.

The authors frequently draw their experience of working with teachers and an interesting example in the chapter on "Teaching and Learning" is a matrix produced by Coventry teachers to record and analyse children's writing. This matrix could also be used to help teachers develop a curriculum for helping children acquire certain skills. There are references throughout the book on how to use analyses and recordings to promote change in practice.

Chapters 6 and 7 deal with relationships and how to cope with those relationships that break down. Thus the emphasis for behaviour problems in this book is framed in terms of presentation and shared responsibility between child, parent and teacher. The authors offer a positive framework for problem solving which can and should at times involve the whole school. This leads neatly to the final chapter on whole school response and policy development and the importance of leadership in the school.

The book is well written, the layout is helpful and aids learning, and the content is broad-based. For me the authors move swiftly from "special needs children" to individual needs and how to respond. In some ways the book is quite charismatic and the feeling I was left with was that if I did all this as a teacher then all would be fine. I know that reality is not so simple but the book offers, I believe, a clear starting point for primary teachers. I would like to see it read and discussed by primary teachers and the content digested and adapted by schools in a creative way, rather than followed slavishly.

I think the authors have been successful in drawing together the main underlying threads in good current educational practice and applying them to individual needs. This book would be useful to all primary teachers, educational advisers, and teachers in training. I will finish with a quote from Peter Robinson in the foreword: "Their (the authors') essential argument is that if we can get the school functioning well, everybody benefits — parents, teaching and non-teaching staff, the head and, most important, *all* the children."

TERI-ANNE HORNBY

COLLEY, A. M., AND BEECH, J. R. (Eds.) (1989). *Acquisition and Performance of Cognitive Skills.* Chichester: Wiley, pp. 348, £32.95 hbk., ISBN 0-471-91461-4.

The pen is not actually mightier than the sword but it can be a great leveller of ideas, flattening the landscape with a homogenised flow of words. Nothing demonstrates this better than the recent cultural phenomenon of edited texts made up of tutorial essays. We are usually assured by the series editors that these fill gaps in the current exposition of research. What they seem to have in mind, however, is something analogous to filling in potholes in the road, guaranteeing a smoother ride to the drive-in cinema and crematorium. How much preferable it would be if editors sought out and presented unexpurgated the rough peaks of scholarship, enabling more of us to share the excitement of new ideas.

There is a place for the tutorial essay. It can serve an eminently useful function in providing a simplified overview of inherently complex or technically difficult material; it can integrate radically new ideas from disparate sources; or it can organise and summarise large quantities of data. Such essays appear from time to time as journal articles or separate monographs. In contrast, the essays in edited texts are typically portraits of research topics, having been (a) requested by an editor and (b) written to an arbitrary deadline. Their major use perhaps is as cannon-fodder for undergraduate essays.

The authors in the present book do a diligent job presenting portraits of topics of cognition and skill. Some of them have had enough practice on their portrait to produce it in their sleep. Some have attempted to assimilate and organise recent data and ideas. All too often, though, they are bland exercises presenting material of variable quality. Whatever aspirations the editors had for a coherent text, the portraits sit side by side uneasily nodding to each other, seeking affinity through their use of the blanket term cognition.

The first and longer half of the book deals with acquisition of skills, and the second half with conditions of performance, including individual differences. The editors gracefully set the scene in an introductory chapter and balletically tie up loose ends in an epilogue. It is a civilised occasion.

Two chapters in the first half deal with the acquisition of fairly specific skills, learning to read English (Beech) and learning to write computer programs (Elsom-Cook). Annett attempts to discern general principles of training. Gilhooly and Green discuss the learnability of problem-solving skills and draw inspiration from computer simulations of skill. Colley discusses the acquisition of motor skill in the context of recent process theories of skill. The longest chapter, by Sincoff and Sternberg, attempts a theoretical overview of the development of skill in children, which becomes lost in clouds of specious abstraction.

Three chapters in the second half deal with conditions of performance: they cover respectively ageing (Salthouse), stress (Hartley, Morrison and Arnold) and diurnal rhythm (Smith). The kind of wisdom to be found in these is that one is not at one's mental peak after a heavy lunch, after prolonged loss of sleep, or beyond the age of seventy. Barber, discussing attention and the ability to do two tasks at once, can also be said to deal with special condition of performance. The remaining chapter, by Ericsson and Oliver, describes an exercise in the training and testing of mnemonic memory.

Since this book is no worse than others of its kind and better perhaps than some, it may seem unfair to use it to attack the genre. Let me make it clear that I am criticising the genre, and that this happens to be the book I was asked to review.

HENRY SHAFFER

EISENBERG, N., AND STRAYER, J. (Eds.) (1987). *Empathy and its Development.* Cambridge: Cambridge University Press, pp. x + 406, £32.50 hbk., ISBN 0-521-32609-5.

I think many psychologists will find this a very uncomfortable book indeed. The notion that anyone can (or would want to) research into something as slippery as empathy is likely to send all committed behaviourists screaming back to their laboratories. Empathy, like love and dreams, is very much an inner experience, belonging to the world of subjective experience rather than to that of public information and the hard data of science.

Yet any behaviourist who takes the trouble to read this book must surely pause to re-examine his or her prejudices. As the editors point out in their preface, research into empathy has increased dramatically in recent years, and far from being woolly-minded and redolent of moral uplift most of it lays as much claim to being scientific as the majority of other things we do in psychology. And if justification is needed for a focus upon empathy, the editors provide it when they indicate how important it is in the development and maintenance of affective bonds, of understanding, and of caring actions between people. The experience of empathy is certainly a subjective one, but the *operation* of empathy is very evident in social behaviours, and may indeed be one of the strongest elements in determining the direction and content of these behaviours.

The book makes clear, however, that we still need a firm consensus on what empathy actually is. Some psychologists take it to refer to cognitive process analagous to cognitive role taking or perspective taking, others take it to mean an affective process with cognitive components, while yet others hold it to be a communicatory and/or information gathering function. Even amongst the contributors to this book, no final agreement seems to emerge. However, the editors' own contention is that empathy involves "sharing the perceived emotion of another" — "feeling with" another. As such it can be distinguished from sympathy (feeling *for* another), from personal distress (one's own aversive feelings aroused by the plight of another) and from projection (the act of ascribing to another one's own feelings, attitudes, etc.). And without wanting to put words into the mouths of the rest of the contributors, it seems to me that this definition would raise no major objections from any of them.

If this is what empathy is, the next question to ask is how do we research into it? Space doesn't allow an extensive summary of the research presented in this volume, but among the more frequently mentioned techniques are (predictably) self-report, observer reports on (infants and young children's) reactions to distress in others, picture-story indice scores, questionnaire data, facial/gestural indices, teacher/parental reports, motor mimicry, and somatic or physiological measures. Reading the data generated by these measures, I was struck by the consistency of the findings they appear to generate, and by the fact that where differences exist they seem to be due more to disagreements over definition rather than to inadequacies of methodology.

In addition to summarising research, the contributors also look at the kind of environment in which empathy (in itself an apparently innate response) is most effectively fostered in children. Such an environment is one that satisfies the child's own emotional needs, discourages excessive self-concern, encourages the child to identify and experience a broad range of emotions, and provides numerous opportunities for the child to observe and interact with others. Nothing exceptional about any of these of course. They represent the kind of environment most conducive to the development of most of the major aspects of psychological health. But it is nice to see them spelt out here, and supported by sound research evidence.

In all, therefore, a very welcome addition to the literature. I hope the book commands the attention it deserves.

DAVID FONTANA

GREER, B., AND MULHERN, G. (Eds.) (1989). *New Directions in Mathematics Education.* London: Routledge, pp. 337, £30.00 hbk., ISBN 0-415-00541-8.

The book builds upon a conference held in Northern Ireland on developments in teaching mathematics, and five of its chapters have their origins in this conference. It aims to "examine interdisciplinary aspects of mathematics teaching using a variety of practical and theoretical approaches to learning and instruction" to "view problems in mathematics education from an international perspective" and to "provide a valuable source of debate and ideas for mathematics teachers, psychologists, educational researchers, administrators, school inspectors, and mathematicians with an interest in teaching their subject". Studies relate to children in the age range 4-16 years.

There are four sections, a foreword by Kath Hart and an overview from the editors. The first section considers how psychological insights can contribute to mathematics teaching. Greer's chapter charts the changes which are occurring in both cognitive psychology and in mathematical education. The cognitive revolution against behaviourism has led to the need for stories about mental representations and processes, not just establishing links between input and output. Increasingly, studies of cognition involve studies of personal interactions, past experiences and social context. Mathematics educators have accepted the need to relate mathematics to everyday life; the need to introduce new content in the shape of uses of information technology and statistics; the need to reform assessment methods, and the need to devise curriculums which build on past experience and what pupils can do, rather than on what a few may need in order to cope with university mathematics courses.

Mulhern's chapter addresses the problems of making inferences about mental processes, and offers a brief review of methods of error analysis, protocol analysis, and psychophysical techniques. He points to the individualised nature of children's mathematical thinking, and argues that techniques of clinical interviewing should become an essential part of every teacher's repertoire. De Corte and Verschaffel provide an important and interesting review of studies of LOGO. Few of the claims made for LOGO are supported by systematic study. For example, "results. . . falsify the widespread view that Logo is easy for children" (p. 68); "there is scant evidence. . . that children acquire important and transferable learning and thinking skills" (p. 74); "there is again no strong empirical evidence that . . . Logo has significant positive effects on children's achievements in school mathematics" (p. 78). The authors offer a number of reasons for these failures, one of which is that at least 30 to 40 hours of pupil time may be necessary to show any benefits, supposing that better tutorial strategies could be devised.

The second section focuses on the basic mathematical operations. It begins with another good chapter by De Corte and Verschaffel who discuss the teaching of word problems in primary school. A strength of the chapter is the wide variety of studies which it calls upon, including a detailed analysis of problems, paper and pencil tests, interviews, observation of problem-solving behaviour, teaching experiments, interviews (which include asking pupils to describe the problem), and an analysis of texts. The authors point out that the stereotyped supply of problems leads pupils to develop superficial solution strategies (e.g., picking out keywords like "more" and "less" which can almost invariably be taken as instructions to "add" and "subtract" respectively). They offer good advice to teachers on the need to offer a variety of problem types, to phrase problems carefully, to make problems meaningful, to take account of pupils' informal solution strategies, to look for systematic patterns of error, and to use multiple problem representations. Mangan's chapter addresses multiplication and division, and provides a literature review together with recommendations for the teacher. The empirical base for this advice is not always clear, nor is it obvious exactly how the teacher should proceed in order to prevent misconceptions by better design of early learning experiences, or to extend pupils' mathematical experiences, or to provide rules to counteract misconceptions, for example.

The third part of the book examines two curriculum developments in the UK: one is an A-level Further Maths project carried out in Northern Ireland, which was designed to allow the spirit of Cockcroft to ascend beyond GCSE; the second is a description of some statistical investigations developed by the Open University. A third chapter in this section offers a refreshing surprise in the form of a consumer's view of GCE, CSE, and GCSE. Gerry Hamill teaches mathematics in an inner city school in Belfast. His chapter considers some of the problems posed by different examination questions, most notably the way that unnecessarily complex language compounds question difficulty for many pupils. The chapter raises (implicitly) the problems that examination boards have in trying to establish the difficulty of questions in the uncharted territory of GCSE. Perhaps the time has come for the difficulty of examination questions to be determined by trialling rather than the best guesses of experts.

Part four takes an international view of curriculum development. Elizabeth Oldham asks if there is an international curriculum in mathematics, and bases her answer on the Second International Mathematics Study. She offers a clear description and an interesting analysis of this study. Hoffman considers the changing mathematics curriculum in the US, and points out how little has changed over the last 50 years. The author directs the University of Denver Project which aims to reform the curriculum in (too much?) detail. The style is somewhat polemic in parts, e.g., "in problem solving it [the calculator] is unequalled in power" (p. 249). The title is too grand: the paper can't hope to review the US as a whole. The largest mathematics project in the US (the University of Chicago School Mathematics Project) is discussed only briefly, and then mainly to contrast the author's views on what they should do with what they are actually doing. Lynn offers a chapter on mathematics teaching in Japan in which he examines some of the underpinnings of their undoubted success. This rich and convincing analysis of a wide range of social and organisational factors is well worth reading. It is a pity that the analysis is trivialised in the conclusion which states "the abiding impression about the achievements of the Japanese in mathematics education is that they are not the result of any special organisational or pedagogical method, but mainly attributable to nothing more mysterious than hard work" (p. 283).

The final view from Greer and Mulhern addresses the problem of improving mathematical education, and some ideas on how this might be brought about. They point to past problems of the inaccessibility of research to teachers, and to the new climate where teachers often see themselves as researchers. There are resource problems, and more general problems of overfacing teachers with a plethora of demands for change not only in content, but also in methods of teaching and assessment. The authors suggest a set of priorities for innovators which include: better treatment of teachers, more in-service training; greater emphasis on primary mathematics, and paying greater attention to research. This concluding section, taken with the foreword by Kath Hart, forms a nice bracketing to the book. Hart addresses a number of important issues, such as the perils of adults presuming to know which problems are realistic to children, the need for teachers to explore the problems of learning for themselves, and to be aware of the intellectual leaps they ask pupils to make, the dangers of labelling some pupils "low ability", and offering them an impoverished diet of mathematics as a result. She points to the lack of current research in the UK. Most tellingly, she makes a plea that the claims about effort, enjoyment and enthusiasm from pupils which appear in some research reports be tested out in representative (rather than trial) classrooms before they are proposed as activities to be taken seriously by teachers.

What of the book as a whole? It would be hard to give an overview of the current scene in mathematical education in the space of 300 pages. Basing a book on conference papers adds to the problem. The book's sectioning reveals some of the difficulties, with a section devoted to international curriculum developments, one to UK curriculum developments, and one each to psychological perspectives and the basic arithmetical operations — certainly a broad canvas, but hardly a balanced one in any sense. Inevitably, coverage is patchy within each section, and the quality of the chapters is mixed. Brian Greer and Gerry Mulhern have made a brave attempt to provide a unifying framework for the disparate elements in their book. One does finish reading with a feeling that some cautious progress is being made in the slow process of bringing together teachers, mathematicians, psychologists, and educational researchers to tackle the major problems faced by mathematical education. It is clear that this coming together needs to be accelerated rapidly, given the demands and short time scale of the National Curriculum. The book is acceptably presented, and the editing has been done with care (although the index refers to Judith not Judah Schwartz). The book does provide a "source of debate and ideas", as promised, and is worth reading, but perhaps not £30 for one's own copy.

JIM RIDGWAY

HOLLY, M. L., AND MCLOUGHLIN, C. S. (Eds.). *Perspectives on Teacher Professional Development*. Basingstoke: Falmer Press, pp. 326, £24.00 pbk., ISBN 1-85000-345-9; £10.95, pbk., ISBN 1-85000-346-7.

The increasing demands on teachers in societies which appear to be requiring ever greater returns from schooling have brought into focus a series of questions about their professional competence and development. This collection of papers on the theme is divided into five sections: contexts of professional development, beginning teaching, experienced teachers, support for professional development and new images and directions. Each section contains two or three articles.

There is little that is new in the collection, most contributors synthesising and filtering existing work rather than adding to it, although Holly's study of teachers' own perceptions of professional development, based on interviews with sixty teachers in Britain and the United States, is a welcome exception. She found that most teachers in both countries claimed to have obtained the majority of their insights and ideas about teaching from other teachers, and preferred visits to different classrooms, rather than reading or in-service courses, for their own further professional development. One interesting contrast between the two groups was that American teachers were more likely to have been on in-service courses than their British counterparts, but were less likely to give courses a high value rating.

Some authors offer models, though, as someone who has seen more than his fair share of concentric circles, Venn diagrams and flow charts, I often wish these were less neat and more craggy, like life itself. In a chapter on adult ages and stages Oja reviews the perspectives of life age/cycle and cognitive-developmental stage theorists. Some of the conceptualisations she describes seem reluctant to assign any labels to life beyond the age of 50 or 55, which must be unsettling for teachers with ten or fifteen years of their professional career in hand and no classification to claim.

Oja concentrates on Loevinger's ego developmental stages which range over self-protective, conformist, self-aware, conscientious, individualistic, autonomous and integrated ego phases. She illustrates this with case studies of teachers who seem to fall remarkably well into their assigned category. Thirty-four year old Elliot, for example, is at the "settling down after 30s transition, individualistic level" when, according to the author, "having dealt with the issues of the Age 30 Crisis, men settle down in the early thirties, reaffirming commitments and often choosing the career as the most highly valued investment of time and energy". One wonders how real these reified concepts are. Presumably if we had a duodecimal counting system Elliot would not have had his crisis until he was thirty-six. I am sure these sequential schemata are more overlapping and interlocked than their proposers claim, and the more superficial have in any case been drawn up from studies of middle-class white males.

In an interesting chapter on beginning teachers Tickle explores the ways in which initial training can avoid the dependence on coping strategies that can easily come from a pure apprenticeship model. He sees early professional development as a time when new recruits can learn to analyse children's work, appraise the processes in their own classroom, and base their practice on something that is empirically devised. In another chapter in the same section Appelgate explores the notion of "readiness" to teach amongst newcomers, likening it too easily to reading "readiness".

This is a useful if varied collection of pieces on professional development and the chapter by Holly and McLoughlin on journal writing for teachers is one of the few articles I have seen on something that is often advocated as a means of developing the reflective practitioner, sometimes employed, but rarely studied. As ever with Falmer Press, the book is beautifully produced and easy to read and handle.

E. C. Wragg

Johnstone, M. (1989). *Stress in Teaching: An Overview of Research.* Practitioner Minipaper 5. Edinburgh: Scottish Council for Research in Education, pp. 51, £4.20, pbk., ISBN 0-947833-29-3.

Fontana, D. (1989). *Managing Stress.* London: BPS and Routledge, pp. 114, £5.95, pbk., ISBN 0-901715-97-2.

When I first started researching the topic of teaching stress in 1975, there was very little published work to refer to. In contrast, in 1989 the National Foundation for Educational Research (Rees, 1989), the Northern Ireland Council for Educational Research (D'Arcy, 1989) and the Scottish Council for research in Education (Johnstone, 1989) have coincidentally published studies commissioned to review the existing literature on this topic and, in addition, Cole and Walker (1989) have published a book based on papers presented at a National Conference on this topic held in 1987. At the same time, a number of books have appeared aimed primarily at helping teachers and/or others to cope with stress. I once started a draft of a paper saying that "since I began researching teacher stress, stress in teaching has risen", which I altered when a colleague point out that it might imply I was claiming to be the cause! Nevertheless, I think I can safely say now that teacher stress seems to be particularly acute at present, and is certainly a prominent item on the educational agenda.

As such, the books by Johnstone and Fontana are timely and welcome. Johnstone's book is the latest in a series of minipapers published by SCRE. This series attempts to be of use to teachers by being both scholarly (in drawing upon research) and clear (in providing readable and untechnical prose). Johnstone's book displays these two qualities admirably in presenting an excellent overview of this topic. The main chapters of the book address six key questions: what is stress?, what are the causes of stress?, how prevalent is stress in teaching?, who is stressed?, what are the effects of stress?, and how can teachers cope? There is a useful summary note at the beginning of each key chapter and a final summary chapter, a technique that will be helpful to meet the needs of those readers who wish to skip read through the book and well as those wanting to take a more considered and reflective approach.

Johnstone notes the problems in trying to define stress but does not offer her own view. Indeed, the whole book is more a commentary rather than a theoretical or conceptual contribution to the topic. In a sense it can be regarded as a well considered briefing report for readers, and as such Johnstone did well to remain somewhat detached from some of the key debates highlighted. As an informative and readable guide, this book can be highly recommended.

Fontana's book is also part of a series aiming to bring research findings to a wider audience of practitioners in a helpful manner. The series, called "Problems in Practice", focuses on a problem from a psychological perspective that is common to a number of professions. Stress is certainly one of these although, interestingly, little work has been done comparing occupational stress across professions, even though a number of writers have conducted studies with different professions. Fontana's key questions, used to shape the book, are: what is stress?, how stressed are you?, what causes too much stress?, and managing yourself. Fontana manages, quite remarkably, to write an informative overview of a substantial length without citing references. Given the reflex action to do so that almost all academics have built up over the years, this is no mean feat, and I suspect a number of writers would simply not have been able to do this, let alone do it well. Clearly, there is a real need for books of this type, that can be read easily cover to cover, and incorporate our current state of the art wisdom on the topic rather than describe and interpret research findings. Fontana does this well, but also does make his own contribution where appropriate. For example, in discussing definitions of stress, he does offer his own: "Stress is a demand made upon the adaptive capacities of the mind and body."

The style of the book is explicitly that of assuming the reader has a problem and the book is offering advice and guidance — in a sense the book is an extended counselling session. "You" and "your" are used frequently, rather than the detached prose normally used in overviews. Plenty of examples are also provided to engage the reader and embed the points made in real-life situations. This is a strength of the book in many respects, but also a drawback in as much as it typifies experiences through examples which underemphasise how varied individuals' responses to a situation are. It also, to some extent, masks the genuine difficulty in offering an interpretation of the situation described. Fontana offers his, but other writers could offer equally plausible different ones.

Because the book is aimed at a range of professions, teachers will find many examples are given from outside the world of teaching. In some ways this can be very helpful, particularly in conveying that the problem of stress is not peculiar to the nature of teachers and teaching. Whilst clearly this is not the case, most books on teacher stress do convey a sense of introspection into the world of teachers and their problems. In thinking about stress, a more distanced view can be refreshing. At the same time, however, it does have the shortcoming of not focusing on the nature of stress experienced by teachers explored in the context of the world of schools and how they operate. There are many very specific aspects of the demands made on teachers that can be usefully highlighted, but Fontana is restricted in being able to do this because of the general audience the book will have. I also think that the omission of a list of further reading was probably a mistake, since I suspect a number of readers will want to follow up their reading with reading specific to their profession. Three or four readings for each of the main professions towards whom the book is directed would have been a useful addition.

Nevertheless, as an informative and readable guide to stress that aims to be of practical help to a range of professions, including teachers (and dare one say nowadays, including teachers in further and higher education?), this book can be highly recommended.

CHRIS KYRIACOU

JOYNSON, R. B. (1989), *The Burt Affair*. London: Routledge, pp. 347, £24.00, hbk., ISBN 0-415-01039-X.

Sir Cyril Lodowic Burt (1883-1971) was the founding father of British educational psychology and was actively associated with this Journal from its inception in 1923 as the *Forum of Education*, through its change in title in 1931 and until his death. During most of this time the Journal was edited by Burt's close friend, C. W. Valentine, and although there were changes on the editorial board, Burt remained on it until his end.

After graduating at Oxford in 1906 and serving from 1908 as assistant lecturer in physiology and lecturer in experimental psychology at the University of Liverpool, he was appointed in 1912 to the post of psychologist to the London County Council, moving to the capital in 1913. His duties were initially concerned with the examination of pupils in elementary schools nominated for admission to schools for the mentally deficient, and he was a member of the Inspectorate, rather than of the medical department. In order to do his job competently he needed tests, so set about translating and standardising the Binet tests and inventing group tests of various kinds. His research interests were prodigious, including the development of factor analysis, and articles, reviews, books and reports soon poured forth from his pen, written in stylish immaculate English.

In 1932 he was appointed to the most prestigious chair in psychology in the land at University College, London, succeeding Charles Spearman, and for many years was regarded as the most distinguished British psychologist, once described as a polymath of renaissance dimensions. He retired in 1950, but continued writing and publishing widely quoted articles and kept up an active correspondence with scholars throughout the world, almost until the time of his death twenty-one years later. One of his enduring interests, which embraced both the subnormal and the gifted, was the question of the role of heredity in human differences, particularly with regard to IQ and scholastic achievement, and some of his seminal papers on this topic were published in this Journal, while others appears under his own or others' names in the *British Journal of Statistical Psychology* of which he was an editor.

Leslie Hearnshaw, professor of psychology at the University of Liverpool, and noted historian of British psychology, was invited by Burt's sister to deliver the address at a memorial service, and so pleased was she with his oration that she asked him to write the official biography of her brother, making available letters and diaries in addition to the mass of published material. This involved an enormous amount of research on Hearnshaw's part, and, shortly after he started writing, there appeared an article in the *Sunday Times* (24th October 1976) in which Oliver Gillie, backed by ourselves and a colleague accused Burt of fraud.

During 1974 three publications had cast doubt on the validity of his evidence. Kamin (1974), an hereditarian, and ourselves, occupying an intermediate position (Clarke and Clarke, 1974) had independently drawn attention to serious problems, particularly with the studies of identical twins, reared together and apart and the question of intelligence and social mobility. The discovery by Gillie that two of Burt's collaborators, who had been responsible for much of the testing and who had published articles from University College, London, in the 1950s, could not be identified or located convinced us that something was very seriously amiss.

Hearnshaw's biography appeared in 1979 after further additional hard and exacting research had been undertaken to explore these and other charges. His book, which the present reviewers heartily recommend, presents a very cautious evaluation of an enormously complex character, admired by many, feared by some, and regarded as devious and corrupt by others. His main conclusions were that Burt had pioneered a number of important areas of research, but that during the post-war period "the verdict must be . . . that at any rate in three instances, beyond reasonable doubt, Burt was guilty of deception. He falsified the early history of factor analysis . . .; he produced spurious data on MZ twins; and he fabricated figures on declining levels of scholastic achievement. Moreover, other material on kinship correlations is distinctly suspect" (p. 259). Hearnshaw at the same time drew attention to Burt's many scholarly and practical achievements.

Ten years after the publication of Hearnshaw's biography, which had laid to rest the bitter controversy following Gillie's article, Joynson has reassessed the evidence in a lengthy book.

Starting from a position of little interest in the Burt affair, he did not read the biography until 1980. The initial impetus for his ultimate rejection of Hearnshaw's conclusions arose from a check he made on the early history of research on intelligence. Hearnshaw had indicated that Burt had given false accounts of his own contribution, diminishing the claims of Spearman. To Joynson, Burt's account seemed correct and thus "it was unthinkable that I should not go on to examine the further allegations against him. . . . So I found myself involved, quite unexpectedly, in an examination of the whole case against Burt. I have come to the conclusion that the evidence which has so far been presented is insufficient to support the accusations which have been made. . . . A grave injustice has been done" (pp. viii-ix).

Joynson is, however, sufficiently sensitive to the evidence not to present Burt as a paragon of virtue, but he takes issue at considerable length with the major criticisms against him. To achieve this he unfortunately uses every available means, including gossip and innuendo, to cast doubt upon the integrity and scholarship of Burt's critics, particularly Leslie Hearnshaw and ourselves. It would be inappropriate to refute here the many arguments advanced against those associated with the attack on Burt. Suffice it to say that in the areas with which we are most familiar, including the thesis summaries written by Burt and published by him in our name in this Journal in 1950, he has made many errors of fact. It may be that in his zeal to vindicate Burt, Joynson has been less than scrupulous in his search for, and assessment of, the considerable documentary evidence. A more important error is that he paraphrases our statement "Burt probably had little if any systematically collected MZA data" (Clarke and Clarke, 1980a and 1980b) as "the Clarkes have expressed their view that Burt never collected kinship data" (Joynson, p. 165). Joynson accuses Hearnshaw of bias but the charge could as easily be reversed. It is a pity that academic controversies become so polarised, but perhaps this is inevitable when the issue at stake, scientific fraud, is so serious.

So far as the augmentation of twin data in the 1950s is concerned, and particularly the activities of Miss Conway who is said to have tested those reared separately as well as writing articles in her own name, Joynson offers an alternative account which seems to us very unlikely, and in any case does not absolve Burt of gross deception of trusting scholars all over the world; not only were these quoting his work, but were also writing to him for further details. Miss Conway remains unaccounted for by those who knew Burt best, nor is there any reference to her in his diaries or in any London records. Yet, among other things "she" was able to write in Burt's journal about the social class of London school children "tested by Burt and myself in 1950" (Conway, 1959, p. 11). This article was, typically, written by Burt himself!

There remain, and presumably always will remain, many puzzles about Burt as a psychologist and as a man. His work on heritability has been superseded years ago, and the nature-nurture debate continues. His books on backward children, young delinquents and the gifted continue to give pleasure to some, but they cannot fail to become progressively dated. Historians of psychology may well continue to be curious concerning the Burt affair, and to them we would say the matter is extremely complex, and judgment should rest on a detailed knowledge of many documents in the Burt archives in Liverpool as well as a thorough acquaintance with Hearnshaw (1979) and papers written by Burt or under one of his pseudonyms. Joynson's (1989) book may be difficult to interpret accurately without this background knowledge.

ANN AND ALAN CLARKE

REFERENCES

CLARKE, A. M., AND CLARKE, A. D. B. (Eds.) (1974). *Mental Deficiency: The Changing Outlook.* 3rd edn., pp. 168-171. London: Methuen; New York: Free Press.

CLARKE, A. M., AND CLARKE, A. D. B. (1980a). Comments on Professor Hearnshaw's "Balance sheet on Burt". In Beloff, H. (Ed.) *A Balance Sheet on Burt.* Supplement to the *Bull. Br. psychol. Soc.*, 33, 1-8.

CLARKE, A. M., AND CLARKE, A. D. B. (1980b). Review of *Cyril Burt: Psychologist* by L. S. Hearnshaw. *Br. J. Psychol.*, 71, 172-173.

CONWAY, J. (1959). Class differences in general intelligence: II. *Br. J. stat. Psychol.*, 12, 5-14.

HEARNSHAW, L. S. (1979). *Cyril Burt: Psychologist.* London: Hodder and Stoughton.

JENSEN, A. R. (1974). Kinship correlations reported by Sir Cyril Burt. *Behav. Genet.*, 4, 1-28.

KAMIN, L. J. (1974) *The Science and Politics of IQ.* Potomac, Md.: Erlbaum.

RICHARDSON, V., CASANOVA, U., PLACIER, P., AND GUILFOYLE, K. (1989). *School Children At-Risk.* London, New York, Philadelphia: The Falmer Press, pp. 286, £9.95 pbk., ISBN 1-85000-515-X; £25 hbk., ISBN 1-85000-514-1.

WEHLAGE, G. G., RUTTER, R. A., SMITH, G. A., LESKO, N., AND FERNANDEZ, R. R. (1989) *Reducing the Risk: Schools as Communities of Support.* London, New York, Philadelphia: The Falmer Press, pp. 282, £9.95 pbk., ISBN 1-85000-531-1; £18.95 hbk., ISBN 1-85000-531-1.

Both books attend to the same broad theme: children "at risk". Both have their merits. Neither seems to me to be entirely appropriate for the UK scene. Of the two, *School Children At-Risk* seems to have the most to offer a UK readership. This is because, of the two books, it is the one which is able to treat the term "at-risk" problematically and theoretically, and because it provides a more ethnographic approach to the problem.

The major problem I have with Wehlage *et al.* is its unproblematic reification of the term "at-risk youth". It is the kind of problem that all authors and researchers in this kind of area face: how to write and research the issues without labelling and categorising young people in ways which turn them into problems, and in so doing, let the social systems, institutions which "process" them off the hook? The dilemma reveals itself in such statements as, "All too often school counsellors have neither the time nor training to become advocates for young people, nor to help students recognise their problems and avail themselves of community resources in response" (p. 23). The issue, for me, is about the taken-for-granted student "dependency", "passivity" and the student as having problems which is everywhere in the book. There is no where that the alternative of "student as advocator", or "student as agent of change", or "student as generator of the agenda" is to be seen. One cannot doubt the examples of "good works" and good intentions that abound in the book. But nowhere is there that critical edge which could enable and empower people to critique the social circumstances that brought them to their situation in life. Has it been so long ago that people like Paulo Freire wrote, or Herbert Kohl or Carl Rogers? Or, is it that they are considered irrelevant to the issues surrounding the notion of "at-risk"? They could certainly provide the book with a more critical edge.

School Children At-Risk at least goes some way towards providing a critical insight into the notion of "at-risk". They offer two models: the epidemiological and the social constructivist. They reject the former and in accepting the latter consider that "the person considered at-risk, the reasons for this consideration, and the ways the school responds are viewed as being constructed in the context of the classroom" (p. 7). They go on to recognise that the labelling involved "could be more damaging to a child than the characteristics of the child that led to the designation" (p. 9). They do not, however, entirely escape the language through which it is the child who becomes identified as the problem and not the institutions and practices of the social context.

In general, both books appear to me to have little to offer a UK readership. But of the two, I would recommend *School Children At-Risk.*

JOHN SCHOSTAK

STEVENSON, R. (1988). *Models of Language Development.* Milton Keynes: Open University Press, pp. 125, £8.95, pbk., ISBN 0-335-09520-8.

BENNETT-KASTOR, T. (1988). *Analysing Children's Language.* Oxford: Basil Blackwell, pp. 149, £9.95 pbk., ISBN 0-631-16375-1.

GARTON, A., PRATT, C. (1989). *Learning to be Literate.* Oxford: Basil Blackwell, pp. 249, £8.95 pbk., ISBN 0-631-15835-9.

The development of children's abilities to use language, while a natural topic of perennial interest among linguists, psychologists and, to a degree, teachers, has undergone two periods of more intensive debate. The first, occasioned by Chomsky's fierce reaction to behaviourist accounts of the language learning process, led to a significant reshaping of the theoretical map, and a flurry of research and polemic. The second, more recently stimulated by the naturalistic studies of researchers such as Gordon Wells, has also added a new dimension to theory, in the taking into account of social interaction as an explanatory context. A new wave of research and writing is at present under way and the three books reviewed here are all products of this recent increased interest. All three attempts, in different ways, to review "state of the art" knowledge about various aspects of language development, and all three will find a ready readership amongst differing groups.

In *Models of Language Development* Rosemary Stevenson presents a treatment of some of the major issues in language development. The book is aimed primarily at students of cognitive psychology and has the particular, useful feature of addressing its topic from an interdisciplinary standpoint. Thus recent developments in linguistics are dealt with in the context of insights from the study of social knowledge. In many ways this treatment puts linguistic flesh on Gordon Wells' process bones. Stevenson's discussion of government binding theory is particularly useful.

The book begins by outlining the sources of knowledge which are integrated in the use of language. The applications of syntax, semantics, pragmatics and social knowledge are illustrated and the book's major thesis outlined: that the development of syntax needs to be considered independently of non-linguistic aspects of language use. The book is thereafter very syntactically based. This inevitably means quite a technical treatment and non-linguists, like myself, will probably find the book hard going. It is a little disconcerting, therefore, to be told, at the beginning of the final chapter, that "this chapter is more technical than the previous ones"! The book does, however, reward careful study, although, as a gesture towards reader-friendliness, a concluding chapter which attempted to summarise the material covered would have been very welcome.

Analysing Children's Language is aimed at a different readership, being more specifically written for child language researchers and students. Its style is, nevertheless, less technical, while maintaining rigour, and much more accessible to non-specialists. This reviewer found it very readable but frustrating — for completely unfair reasons. It is a book which aims to discuss and critically evaluate *methodology*, which it does very comprehensively and interestingly. As one reads the discussion of methodology, however, one continually wants to know, not just *how* research was carried out, but what its findings were! Of course, the fact that the book creates this urge to know things which it makes no claims to deal with may be a very positive attribute. Reading it certainly left this reader in a state of looking for further information, which is compliment enough in itself.

For a reviewer whose chief professional concern is more with the practical implications of theories of language development than with the niceties of the theories themselves, *Learning to be Literate* was a very stimulating read indeed. The book begins with the statement of its assumption that literacy includes the mastery of spoken language as well as the more commonly included skills of reading and writing. This inclusion, while welcome in that it opens up many useful areas of insight, does perhaps require more justification than it is given here. Whether or not the acquisition of spoken language parallels that of written language has itself been the topic of much recent debate, including contributions from such people as Margaret Donaldson (1989). The view one takes on this issue influences one's approach to literacy learning and teaching so fundamentally that a book which *assumes* (rather than argues) one standpoint is at risk of alienating a large portion of its potential readers in the first few pages.

This concern notwithstanding, the book provides an excellent coverage of the field of language and literacy learning. It is up to date in the material it draws upon and insightful in its handling of this. While one could perhaps have wished for a greater proportional treatment of the learning of writing and reading, these chapters in themselves are first rate. There is also a very welcome chapter dealing with metalinguistic awareness, surely an area in line for considerable further research and theory.

Overall the book will certainly figure on lists of recommended reading in many colleges and universities, both in courses of teacher education and of the psychology of education, and will have a justifiable influence.

DAVID WRAY

REFERENCE

DONALDSON, M. (1989). *Sense and Sensibility.* Reading and Language Information Centre, University of Reading.

THARP, R. G., AND GALLIMORE, R. (1988). *Rousing Minds to Life: Teaching, Learning, and Schooling in Social Context.* Cambridge: Cambridge University Press, pp. 317, £25.00 hbk., ISBN 0-521-36234-2.

The dust jacket of this book carries three reviewers' phrases of high praise, thus: "a powerful synthesis of theoretical advances", "a unique contribution", "a major contribution", "this book should be required reading for every professional who teaches", "a model of research and practice", "one of the most exciting books in educational and developmental psychology in recent years". I therefore read it in a sceptical frame of mind. Some considerable reservations remain after reading it, but it is certainly worth serious attention.

The book is multifaced. The central argument is that American education is poor; that current efforts to improve it do not address the central problem and therefore will (like earlier efforts) fail; that the central problem on which reform must focus is that schools run in ways which make teaching unlikely if not impossible, with an underlying theory that students are to learn on their own; that there is an alternative model of what teaching/learning should be emerging from "neo-Vygotskianism"; and that teaching/learning compatible with this model is what should be set up in schools. The Kamehameha Elementary Education Program (KEEP) is presented as "a major 'good example' for teaching, classroom and school organisation, teaching training, and research and development".

The first part of the argument concerns the present awfulness of American schools. They are described as relying on "recitation" in which the teacher asks a series of unrelated questions requiring convergent factual answers and student display of known information. "More effective" teachers manage instructional materials and activities in such a way that students can complete learning tasks with little teacher intervention. Socratic dialogue is conspicuously absent: some nasty incidents of pupil rebellion, teacher stress, general mayhem and unskilled graduates from school are conspicuously present. A further caustic picture of teacher training is presented. I don't feel competent to judge the accuracy of this description of "teaching" being absent from education but readers more involved in schools may find it recognisable. Similar points have certainly been made in British observations of classrooms, and the low level of educational standards has of course been a perennial complaint.

The second part of the argument presents neo-Vygotskian theory as the source of a different way of "teaching". Vygotsky's ideas about the development of verbal cognition, the zone of proximal development, activity settings, cognitive structuring, assisted performance, schooling, and so forth are woven with other researchers' work on language development, scaffolding, and cognitive processes. This section of the book will be a useful introduction to a currently thriving area of developmental psychology. As two of the reviewers I mentioned initially are among the translators and editors of Vygotsky, Tharp and Gallimore's account may be taken to be acceptable to the experts, and it is probably more accessible to non-experts than most writing on Vygotsky as it is clear, lively, and comparatively free from jargon.

The third part of the argument is the embodiment of a better neo-Vygotskian way of teaching in the work of KEEP, and its amelioration of the educational problems of children who would otherwise have done badly in schools. The KEEP educational system developed from a "research-and-development program for improving the cognitive and educational development of a group of educationally at-risk ethnic-minority children". Much of its activity is now being curtailed as "too expensive", "in the face of continuing excellent evaluations of program effects". It is one of the regrettable points about this book that little information is given about program effects: a very substantial proportion of the evaluatory references given are published by KEEP itself and so are comparatively inaccessible, and publications in refereed journals are exceptional. The validity and effectiveness of KEEP methods therefore rest largely on the examples presented. These bear a not inconsiderable resemblance to what has long been regarded as good practice in British primary schools. Tharp and Gallimore do, however, interweave their description of practice with theoretical points which would make it perhaps more useful for a practitioner's reflection than it is for a non-practitioner primarily interested in theory.

I shall add this book to my reading list for teachers doing in-service courses on cognitive development, largely for the sake of its exposition of neo-Vygotsksianism, which is enlivening. People concerned with practice in education, particularly in the primary years, will also find it useful. Its model of what teaching and learning should be, and could be, in schools is sympathetic. It is surprising therefore that the book carried a catchpenny and patronising title: were the children's minds really not alive before KEEP came along to rouse them?

Sara Meadows

Wang, M. C., Reynolds, M. C., and Walberg, H. J. (Eds.) (1987). Oxford: Pergamon Press, pp. 1156, £140.00 hbk.

According to the editors the aim of these three volumes is to give through 45 chapters, a "state of the art" review of research in special education and its links with practice in as many areas as possible. These volumes certainly present valuable summaries of American research in a number of areas of special education and add some interesting dimensions to the debate on how to most effectively support children with special needs.

In trying to cover so much ground inevitably there is not always space to explore the conceptual issues and their influence on research and practice in as much depth as would be desirable, so, for example, limited English proficient children has just over two pages. This raises the question of whether justice can be done to the whole of special education in three volumes given that the American definition of special education also covers children from poor socio-economic backgrounds and ethnic minorities. This problem is exacerbated by the loose framework of the book which misses out whole areas of importance and doesn't have, for example, even one chapter which examines special education from the perspective of the school and its curriculum and organisation. Instead the major focus is still on individual learners and programmes. A collection of papers rather than a handbook of special education would be a fairer description of these three volumes.

Given the almost exclusively American focus on research and practice and the lack of extension on British thinking and development on special needs, plus the price of £120 for the three volumes, these are books which we might want to refer to occasionally rather than own!

Volume 1. Learner Characteristics and Adaptive Education

Volume 1 is divided into three main sections. According to the editors the first section concentrates on the "learning characteristics of handicapped students and the provision of effective education". The second section looks at the various educational programming approaches and the third section focuses on "non-categorical programming for mildly handicapped students". According to Reynolds in his introduction to this section, non-categorical as the name implies is simply a move away from traditional categorisation or labelling and its attendant problems.

Although Reynolds and later authors in this section seem a little unclear in their thinking on this issue some similarities with arguments put forward in the Warnock Report are apparent. Where it perhaps differs is that rather than conceptualising the whole of special education as on some sort of continuum they argue that children with relatively mild handicaps who fall under differing labels such as emotionally disturbed, learning disabled and educable mentally retarded, might be better served by a non-categorical approach.

As is indicated by some of the terminology used, conceptualisation of special needs and its essentially interactive nature seems to be at an emergent stage. According to a paper by Morsink, Thomas and Smith-Davies the adoption of a non-categorical approach is linked to the movement of children into mainstream settings under the principle of the least restrictive environment. The question arising from this is whether more effective teaching and learning is possible with a non-categorical approach as opposed to a categorical one. The attendant issue of teacher training is taken up in the paper by Blackhurst, Bott and Cross. This again shows many parallels with recent thinking in Britain and the move towards a generic training in special needs for those teachers who will be working in a mainstream setting.

Some of the strongest chapters are Michael Gerber's on cognitive behavioural approaches; Margaret Wang's on the role of the learner; and James Ysseldyke's on the classification of handicapped students.

Volume 2. Mildly Handicapped Conditions

Section I claims to deal with mild mental retardation, Section II with behavioural disorders and Section III with learning disability. The first chapter by Polloway and Smith on the mild mental retardation construct confirms the differences between the American and British approaches to conceptualising special needs. Although concern is expressed with labelling, a lot of it is focused on correctly identifying children and redefining the parameters for a particular label. Nowhere in this first paper is the concept of special needs mentioned although the authors do wonder whether it is possible to have labels based on the services required.

Barbara Keogh in her introduction to Section III and her paper on learning disability takes up the issue of how it is conceptualised and how it overlaps with other labels. Gresham in his paper on social competence and learning disabled students gives a very good review of research and intervention in this field. Similarly Lloyd gives a useful if basic overview of "Direct Academic Interventions in Learning Disabilities".

Volume 3. Low Incidence Conditions

This is conceptually a strange collection of papers. Section I deals with the education of deaf children, Section II with the education of visually impaired children and Section III with handicapped infants. What's surprising is that no other relatively low incidence conditions like physical impairment are included. This means that nowhere in the three volumes does physical or motor impairment have any specific import. Given the title of the volume, the indecision of the third section on handicapped infants seems even odder given that in total they hardly comprise a low incidence.

There are five chapters on deaf children starting with Ruigley and Paul's chapters on English language development. They like other authors point out the lack of empirical evidence on approaches to teaching deaf children. Greenberg and Kusche give a good overview of their chapter on cognitive, personal and social development of deaf children.

In the second section on visual impairment David Warren gives a good summary on the development of visually impaired children and attendant research in this area. He stresses the need for sufficient funding for research into such a heterogeneous and widely dispersed population of children and also the need for research in this area to have a functional emphasis. Apart from a chapter on modern technology none of the chapters focuses directly on the education of visually impaired children.

The last section on handicapped infants comprises five useful papers which evaluate carefully the present state of research in areas like assessment and intervention and suggest what research needs to be carried out.

BARBARA RIDDICK

PUBLICATIONS RECEIVED

Books for review have been selected from the following
("n.p." means "no price mentioned")

Ainscow, M. (1989). *Special Education in Change*. London: David Fulton, pp. 149, £9.95 pbk., ISBN 1-85346-127-X.

Ainscow, M., and Florek, A. (1989). *SEN: Towards a Whole School Approach*. London: David Fulton, pp. 158, £9.95 pbk., ISBN 1-8534-126-1.

Arron, P. G. (1989). *Dyslexia and Hyperlexia*. The Netherlands: Kluwer Academic Publishers, pp. 302, £42.00 hbk., ISBN 1-55608-079-4.

Atkinson, D. (1989) *Someone to Turn To*. Kidderminster, Worcs.: British Institute of Mental Handicap, pp. 104, £6.95 pbk., ISBN 0-906054-66-4.

Coleman, A., and Beaumont, J. (1989). *Psychology Survey 7*. London: Routledge, pp. 278, £17.95 hbk., ISBN 1-85433-0012-08.

Felce, D. (1989). *The Andover Project*. Kidderminster, Worcs.: British Institute of Mental Handicap, pp. 56, £5.95 pbk., ISBN 0-906054-70-2.

Hargreaves, D. (1989). *Children and the Arts*. Milton Keynes: Open University Press, pp. 164, £9.95 pbk., ISBN 0-335-09880-0.

Howe, M. J. A. (1989). *Fragments of Genius*. London: Routledge, pp. 178, £25.00 hbk., ISBN 0-415-00781-X.

Jones, N. (1989). *Management and Pupil Behaviour*. Basingstoke: Falmer Press, pp. 277, £9.95 pbk., ISBN 1-85000-592-3.

Jones, N. (1989). *Special Educational Needs (Vol. 2)*. Basingstoke: Falmer Press, pp. 219, £9.95 pbk., ISBN 1-85000-491-9.

Langford, P. (1989). *Children's Thinking and Learning in the Elementary School*. Lancaster, PA.: Technomic Publishing, pp. 171, n.p. hbk., I.S.B.N. 87762-604-9.

Lawrence, J., Steed, D., and Young, P. (1989). *Disruptive Children—Disruptive Schools?* London: Routledge, pp. 260, £10.95 pbk., ISBN 0-415-04676-9.

Lomax, P. (1989). *The Management of Change*. Clevedon, Avon: Multilingual Matters Ltd., pp. 183, £8.95 pbk., ISBN 1-8539-060-6.

MacLeod, F. (1989). *Parents and Schools: The Contemporary Challenge*. Basingstoke: Falmer Press, pp. 233, £9.95 pbk., ISBN 1-85000-499-4.

Makin, P., Cooper, C., and Cox, C. (1989). *Managing People at Work*. London: Routledge, pp. 213, n.p. pbk., ISBN 1-85433-010-1.

McManus, M. (1989). *Troublesome Behaviour in the Classroom*. London: Routledge, pp. 171, £9.95 pbk., ISBN 0-415-03911-8.

Meredith, P. (1989). *Sex Education*. London: Routledge, pp. 181, £25.00 hbk., ISBN 0-415-00604-X.

Postic, M. (1989). *L'Imaginaire dans La Relation Pedagogique*. Paris: Presses Universitaries de France, pp. 161, n.p. pbk., ISBN 2-13-0422543.

Rust, J., and Golombok, S. (1989). *Modern Psychometrics*. London: Routledge, pp. 182, £9.95 pbk., ISBN 0-415-03059-5.

Roland, E., and Munthe, E. (1989). *Bullying*. London: David Fulton, pp. 151, n.p. pbk., ISBN 1-85346-133-4.

Satterly, D. (1989). *Assessment in Schools (2nd edit.)*. Oxford: Basil Blackwell, pp. 354, £8.95 pbk., ISBN 0-631-16982-2.

Steffy, B. (1989). *Career Stages of Classroom Teachers*. Lancaster, PA.: Technomic Publishing, pp. 213, n.p. hbk., ISBN 27762-637-5.

Valsiner, J. (1989). *Development within Culturally Structured Environments. Vol. 1*. Norwood N.J.: Ablex, pp. 260, £27.95 hbk., ISBN 0-89391-487-8.

Valsiner, J. (1989). *Development within Culturally Structured Environments. Vol. 2*. Norwood, N.J.: Ablex, pp. 305, £33.95 hbk., ISBN 0-89391-588-6.

Wheldall, K., and Merrett, F. (1989). *Positive Teaching in the Secondary School*. London: Paul Chapman, pp. 99, £6.95 pbk., ISBN 1-85396-075-6.

Williams, A. (1989). *Physical Education for the Primary Years*. Basingstoke: Falmer Press, p. 182, £8.95 pbk., ISBN 1-85000-535-4.

Zachariah Boukydis, C. F. (1989), *Research on Support for Parents and Infants in the Postnatal Period*. Norwood, N.J.: Ablex, pp. 265, £32.50 hbk., ISBN 0-89391-333-2.

British Journal of Developmental Psychology

Special issue on 'The child's theory of mind'

Submissions are invited for a special issue of the *British Journal of Developmental Psychology* on 'The child's theory of mind'. Empirical and theoretical contributions concerning the development of the child's knowledge of her own mind and the minds of other people are requested. The topic may be approached from any developmental or comparative perspective and may include aspects of emotional development or the developmental psychopathology of metacognition. Papers not chosen for the special issue will be considered for the journal in the usual way. The issue will be published as part of volume 9 (1991).

Papers should be submitted to the Editor, Professor G. E. Butterworth, Journals Office, 13A Church Lane, London N2 8DX, UK *before 28 February 1990.*

The *British Journal of Developmental Psychology* is published four times a year. Subscription price of volume 7 (1989) is £55.00 (US$103)

A special individual price of US$54.00 is available to members of APA/SRCD/ISSBD.

The British Psychological Society
The Distribution Centre, Blackhorse Road, Letchworth,
Herts SG6 1HN, UK.

Journal of Occupational Psychology

Selected contents of recent issues

- Rewards investments alternatives and organizational commitment: Empirical evidence and theoretical development — *N. Oliver*
- Performance evaluation and compensation feedback messages: An integrated model — *B.D. Bannister & D.B. Balkin*
- Coping with job-related stress: The case of teachers — *I.S. Schonfeld*
- Self-ratings, supervisor ratings, position and performance — *P. Herriot & J. Lane*
- Predicting commitment to company and union: Divergent models — *J. Barling, B. Wade & C. Fullagar*
- Consequences of changing employers amongst young engineers — *T.J. Newton & A. Keenan*
- Decision making in the graduate selection interview: A field study — *N. Anderson & V. Shackleton*
- The measurement and antecedents of affective, continuance and normative commitment to the organization — *N.J. Allen & J.P. Meyer*
- Union commitment: An examination of antecedent and outcome factors — *J.W. Thacker, M.W. Fields & L.A. Barclay*

The *Journal of Occupational Psychology* (ISSN 0305-8107) is published four times a year and edited by Dr Michael West (University of Sheffield). The price of volume 63(1990) is £64 (US$119.00).

Members and Foreign Affiliates of the Society and individual members of APA and other selected societies are entitled to purchase journals at preferential rates.

Subscriptions and inquiries to
The British Psychological Society
The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN

Journals of The British Psychological Society

British Journal of Psychology
Edited by A.J. Chapman, University of Leeds
The Journal publishes reports of empirical studies likely to further our understanding of psychology; critical reviews of the literature; and theoretical contributions. Readership is world-wide and includes academic and research workers in experimental and cognitive psychology.
Volume 81 (1990) (4 parts) £97 00 (US$181.00)

British Journal of Medical Psychology
Edited by John Birtchnell, Institute of Psychiatry
The British Journal of Medical Psychology publishes original contributions to knowledge in the area of those aspects of psychology applicable to medicine and related clinical disciplines, including the psychotherapies and counselling. Published articles include empirical studies, theoretical papers and clinical reports which suggest increased understanding of the psychological aspects of medicine.

The readers are medical psychologists, mainly clinicians, concerned with applying psychology to medical problems, from the disciplines of psychology, medicine and professions allied to medicine.
Volume 63 (1990) (4 parts) £76.00 (US$141.00)

British Journal of Social Psychology
Edited by Miles Hewstone, University of Bristol
The British Journal of Social Psychology publishes articles which contribute to the basic methodological and theoretical issues confronting the discipline. In addition, it encourages: theoretical and review papers, papers describing applied social psychology in a variety of settings; correspondence on controversial issues; guest-edited issues on research in growth areas of social psychology
Volume 29 (1990) 4 parts £65.00 (US$121.00)

British Journal of Developmental Psychology
Edited by George Butterworth, Professor of Psychology, University of Stirling
This new journal, launched in 1983, publishes authoritative scientific articles describing all aspects of development from infancy to old age, including educational implications and the problems of abnormal development.
Volume 8 (1990) (4 parts) £61 00 (US$112.00)

British Journal of Mathematical and Statistical Psychology
Edited by A.D. Lovie, University of Liverpool
The Journal publishes articles relating to any areas of psychology which have a greater mathematical or statistical or other formal aspect to their argument than is usually acceptable to other journals. Articles with a clear reference to substantive psychological issues are preferred, e.g. new models for psychological processes, new approaches to existing data, critiques of existing models and improved algorithms for estimating the parameters of a model.
Volume 43 (1990) (2 parts) £64.00 (US$119.00)

Journal of Occupational Psychology
Edited by Michael West, University of Sheffield
An international journal of industrial and organizational psychology which describes and interprets new research into people at work. Areas covered include personnel and vocational psychology; occupational social psychology, human factors and engineering and computer psychology; organizational psychology; behavioural aspects of industrial relations. Innovative or interdisciplinary approaches to problems. are particularly prominent.
Volume 63 (1990) (4 parts) £64.00 (US$119.00)

British Journal of Clinical Psychology
Edited by Ray Hodgson, Department of Clinical Psychology, Whitchurch Hospital, Cardiff
The British Journal of Clinical Psychology publishes new findings, theoretical, methodological and review papers bearing on the whole field of clinical psychology and includes: descriptive and aetiological studies of psychopathology; applications of psychology to medicine and health care; studies of the assessment and treatment of psychological disorders; social and organizational aspects of psychological disorders and ill-health Contributions combine scientific excellence with relevance to the concerns of practising clinicians, bridging the gap between academic and professional interests.
Volume 29 (1990) (4 parts) £65.00 (US$121 00)

All seven journals ordered at the same time for one address £433 (US$805)

The British Psychological Society

The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK.

BEECH TREE SCHOOL

FOR CHILDREN WITH SEVERE CHALLENGING BEHAVIOUR

Beech Tree School offers short planned courses of residential education for children with special learning needs compounded by severe challenging behaviour. The primary goal is to return pupils to their homes and local schools, (some children are prepared for residential schools which can best meet specific learning, sensory of physical needs). A two year course is usually recommended with a period of outreach support at transfer. Both 40 and 52 week residential places are available.

After an assessment period of three months, behavioural, educational and social goals are agreed with parents and the officers of the placing authority. Progress is reviewed fortnightly with a written summary to all concerned every ten weeks.

The school offers formal training to parents and to the staff who will receive pupils when they leave. There is also a home training service for pupils and parents.

During their final months pupils may live in a semi-detached house in the local community to prepare them for life in an ordinary home.

Beech Tree School is DES approved.

Inquiries regarding referrals to - The Head Teacher,
Beech Tree School, Meadow Lane, Bamber Bridge,
Nr PRESTON, Lancashire PR5 8LN. (0772 323131).

A **SPASTICS SOCIETY** EDUCATIONAL SERVICE

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

British Journal of Psychology
British Journal of Medical Psychology
Journal of Occupational Psychology
British Journal of Social Psychology
British Journal of Clinical Psychology
British Journal of Developmental Psychology
British Journal of Mathematical and Statistical Psychology
The British Journal of Educational Psychology

The *British Journal of Psychology*, the *British Journal of Medical Psychology*, *Journal of Occupational Psychology*, the *British Journal of Developmental Psychology*, the *British Journal of Social Psychology* and the *British Journal of Clinical Psychology* appear in quarterly parts. The 1990 subscription prices, payable in advance, are £97.00 (US$181.00) for the *British Journal of Psychology*; £65.00 (US$121.00) for the *British Journal of Social Psychology*; £65.00 (US$121 00) for the *British Journal of Clinical Psychology*; £61.00 (US$112.00) for the *British Journal of Developmental Psychology*; £76.00 (US$141.00) for the *British Journal of Medical Psychology* and £64.00 (US$119.00) for the *Journal of Occupational Psychology*.

The *British Journal of Mathematical and Statistical Psychology* is issued twice a year, in May and November. The subscription rate, payable in advance, is £64.00 per volume post free (US$119.00 overseas).

Subscriptions to any of these journals may be sent to a bookseller or subscription agent, or to The British Psychological Society, The Distribution Centre, Blackhorse Road, Letchworth, Herts., SG6 1HN, UK. A special price of £433.00 (US$805.00) for all seven journals going to the same address is offered for 1990.

The *British Journal of Educational Psychology* is issued three times a year, in February, June and November. The prepaid subscription is £30.00 a year post free, or £10.50 each number. The prepaid subscription price in North America is $60.00 US. Journals are sent accelerated surface post to the USA. Subscriptions should be sent to the Scottish Academic Press, 139 Leith Walk, Edinburgh, EH6 8NS.

MANUSCRIPTS

For the British Journal of Psychology should be sent to Professor A J. Chapman, Journals Office, 13A Church Lane, London, N2 8DX.

For the British Journal of Medical Psychology should be sent to Dr John Birtchnell, Journals Office, 13A Church Lane, London N2 8DX.

For the Journal of Occupational Psychology should be sent to Dr Michael West, Journals Office, 13A Church Lane, London, N2 8DX

For the British Journal of Social Psychology should be sent to Dr Miles Hewstone, Journals Office, 13A Church Lane, London, N2 8DX.

For the British Journal of Clinical Psychology should be sent to Professor R. Hodgson, Journals Office, 13A Church Lane, London, N2 8DX

For the British Journal of Mathematical and Statistical Psychology should be sent to Dr A D Lovie, Journals Office, 13A Church Lane, London, N2 8DX.

For the British Journal of Developmental Psychology should be sent to Professor George Butterworth, Journals Office, 13A Church Lane, London, N2 8DX.

For The British Journal of Educational Psychology should be sent in duplicate to Professor Charles Desforges, University of Exeter, St Luke's, Heavitree Road, Exeter, EX1 2LU.

Contributors are advised to consult the Style Guide issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price £3.50 post free, from The British Psychological Society, St Andrew's House, 48 Princes Road East, Leicester, LE1 7DR

All manuscripts should be submitted in duplicate.

OFFPRINTS

Contributors on application receive fifty copies of their papers free. Additional copies may be obtained in units of fifty copies.

REPRINTS

The Journal is available in reprint from Swets and Zeitlinger B.V., P.O. Box 810, 2160 SZ Lisse, The Netherlands; Vols. 21-29, 31-47.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £40.00.

They are entitled to purchase any or all of the Society's Journals at the special rates of £8.00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester, LE1 7DR.

BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY,

VOL. 60, PART 1, 1990

CONTENTS

ARTICLES:

Eshel, Y., and Kurman, J. — Individualised Instruction, Congruent Authority Structure and Student Outcomes 1

Wishart, J.G., and Duffy, L. — Instability of Performance on Cognitive Tests in Infants and Young Children with Down's Syndrome 10

Hibbett, A., Fogelman, K., and Manor, O. — Occupational Outcomes of Truancy 23

Pierce, C.M.B., and Molloy, G.N. — Pyschological and Biographical Differences between Secondary School Teachers Experiencing High and Low Levels of Burnout 37

Ingram, J., and Worrall, N. — Varieties of Curricular Experience: Backmarkers and Frontrunners in the Primary Classroom 52

Butler, R., and Orion, R. — When Pupils Do Not Understand the Determinants of their Success and Failure in School: Relations between Internal, Teacher and Unknown Perceptions of Control and School Achievement 63

Reid, D. J., and Beveridge, M. — Reading Illustrated Texts: a Microcomputer Based Investigation of Children's Strategies 76

RESEARCH NOTES:

Winter, S. — Teacher Approval and Disapproval in Hongkong Secondary School Classrooms 88

Thompson, S. V., and Riding, R. J. — The Effect of Animated Diagrams on the Understanding of a Mathematical Demonstration in 11- to 14-Year-Old Pupils 93

Seddon, G. M., and Barry, J. F. — The Relative Effectiveness of Four Methods of Teaching Practical Scientific Skills to Secondary School Students . 99

CONFERENCE ABSTRACTS 106

BOOK REVIEWS 116

PUBLICATIONS RECEIVED 129

UK ISSN 0007-0998

AUG 1990

Education

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

JUNE 1990

VOL. 60, PART 2

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by

Professor CHARLES DESFORGES

Assistant Editors:

Dr H. C. M. CARROLL
Dr D. FONTANA
Mr G. HARVARD
Dr P. TOMLINSON
Dr M. YOUNGMAN

Sub-Editor:

Dr S. CUNNINGHAM

Book Review Editor:

FLORA J. MACLEOD

Advised by the Editorial Committee:

Professor G. BROWN
Dr A. MANSTEAD
Professor P. W. ROBINSON
Dr K. WHELDALL

Vol. 60, Part 2

The *British Journal of Educational Psychology* is issued by the British Psychological Society once a term, in February, June and November. The subscription is £30.00 a year ($60.00 North America), or £10.50 each number.

Subscriptions and Orders should be sent to Scottish Academic Press, 139 Leith Walk, Edinburgh, EH6 8NS. Members of the British Psychological Society receive the *Journal* on special terms.

Papers for publication in duplicate should be sent to Professor Charles Desforges, University of Exeter, St Luke's, Heavitree Road, Exeter, EX1 2LU. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for The *British Journal of Educational Psychology* should be sent to: Flora Macleod, University of Exeter, St Luke's, Heavitree Road, Exeter, EX1 2LU.

Advertisements for this *Journal* should be sent to Scottish Academic Press.

Published by
SCOTTISH ACADEMIC PRESS, 139 Leith Walk, Edinburgh, EH6 8NS,
for the British Journal of Educational Psychology, Ltd.

All rights reserved

© 1990. The British Journal of Educational Psychology

PRINTED IN GREAT BRITAIN BY
BELL AND BAIN LTD., GLASGOW

Br. J. educ. Psychol., 60, 133-141, 1990

EXPERIENCES IN SCHOOL AS DETERMINANTS OF BEHAVIOUR IN A DISCUSSION SITUATION

By Einar M. Skaalvik and Olav Sletta
(Department of Education, University of Trondheim, Norway)

Summary. This study examined the effects of school experiences on behaviour in a discussion. Participants were 80 8th grade students (40 females and 40 males). Dyads of students (one high and one low achiever) were assigned to two different solutions of a problem. Each student in the dyad was asked to defend one of these solutions and was given four written arguments in favour of it. The students were instructed to give up the discussion if they were convinced that the partner's position was a better solution. At the end of the discussion the students independently of each other indicated what solution they really thought was the better. Low achievers in school gave up the discussion and claimed to believe that the opposite solution was the better more frequently than did high achievers. There was no significant difference between high and low achievers in the tendency to really believe in the solution they argued for. The results are discussed in terms of self-confidence and self-protection.

INTRODUCTION

The purpose of this study was to examine 8th grade students' behaviour in a discussion and how discussion behaviour is related to school achievement and academic self-concept. For this purpose we defined a discussion as a situation where two students discuss or debate a particular problem that is not part of the school curriculum and that has no right or wrong answer. Appropriate behaviour in discussions is important in democratic processes at all levels. Such processes require that the persons involved both have the psychological strength to stand up for their beliefs and to stay with their beliefs when faced with counter-arguments. Previous research has indicated that poor readers tend to defer to the judgments of good readers, whether or not the issues under discussion are related to competency in reading (Stulac, 1975; Morris, 1977).

Self-concept may partly be seen as expectations of one's own behaviour or performance which affect subsequent behaviour. For instance, some investigators have defined academic self-concept as success expectations in future schoolwork (Brookover and Passalacqua, 1981) or success expectations on specific academic tasks (Skaalvik, in press). In accordance with this view, several longitudinal studies have demonstrated that academic self-concept at time 1 tends to influence academic achievement at time 2 (Shavelson and Bolus, 1982; Marsh, 1987; Skaalvik and Hagtvet, 1990). Thus, self-concept may be expected to influence behaviour.

A critical question is to what situations academic self-concept or self-confidence related to school tasks are generalised. If low or high academic self-confidence is generalised to intellectual areas in general, the students' academic self-concept may influence their behaviour in a variety of situations where intellectual processes are important. In a discussion students with low academic achievement and low academic self-confidence might therefore not expect to be able to argue convincingly about their own position, even if the issue being debated is not part of the school curriculum. These students may also have little confidence that their position is correct or that it is the best, especially when faced with arguments for an alternative position put forward by somebody that they perceive as more able. Through school experiences low achieving students have learned that more able students (in their perception) usually have more acceptable answers or better solutions to problems. Consequently, students with low academic achievement and low self-concept may more often than students with high academic achievement and high self-concept lose faith in the position they hold and change their opinion during a discussion, losing faith that they are able to argue convincingly and hence giving up the discussion.

Academic achievement, social prestige and protection of self-esteem

School is organised in such a way that by and large students of the same age work with the same materials and tasks at the same time and their work is exposed to systematic evaluation. This makes poor achievement an embarrassing school experience that may create a feeling of inferiority. Thus, poor achievement represents a threat to the students' self-esteem and their social prestige. Faced with such threatening situations, low achieving students will often develop self-protective tactics (Rosenberg, 1968; Kaplan, 1980; Breakwell, 1986).

One self-protective strategy would be not to involve oneself or not to put any prestige in the classroom activities. Such a strategy may take different forms or be described in different ways. Rosenberg (1968) proposed that one may try to protect one's self-esteem by lowering the value of those activities where one expects not to do well and putting more value on activities where one expects to succeed better (selectivity of self-values) and by avoiding situations where one expects not to do well (situational selectivity). Covington (1984) suggests that one failure-avoiding tactic used by students may be not to try very hard. Breakwell (1986) claims that one of the targets of coping strategies is to remove aspects of the social context which generate threat.

In the discussion situation not trying or not involving oneself could mean to "surrender"; that is overtly to give up one's position (claim to adopt the other person's position), perhaps without really trying to defend one's own position. This parallels the "avoidance strategy" in defence of self (Kaplan, 1980) and may or may not be followed by a real change of view (cognitive change).

Another self-protective strategy might be just the opposite of surrendering; that is overtly never to give in once one has stated a position, even if one should change one's opinion. This tactic involves a self-presentation that is perceived by the actor to create the most positive image of self. Such attempts to influence other people's perceptions to some extent parallel what Breakwell (1986) in more extreme cases terms "negativism", which involves "attacking the attackers".

Thus, the experience of achieving poorly in school leads to low self-confidence which in a discussion situation predicts loss of confidence in one's standpoints and change to the views advocated by others. The experience of achieving poorly in school also creates a threat to self-esteem and social prestige which predicts either withdrawal from the discussion by overtly claiming to give up one's position (non-involvement or avoidance reaction), or overtly defending one's initial position independently of changes in one's real beliefs (selective self-presentation).

Social self-concept

Social self-concept may be defined as a student's feeling of being accepted by classmates, having good friends and being invited to take part in activities. It is anticipated that students with low social self-concept develop a behaviour intended to increase their social acceptability, or they may try to prevent negative effects of their behaviour on peer acceptance. Hence, they may tend to defer overtly to other students' viewpoints, not because they really accept them, but because non-compliance could result in loss of acceptance. Kelman (1961) suggests that compliance depends on the observability of one's behaviour and the perceived power of the person or persons one interacts with. In a discussion this seems to imply that when faced with a classmate who is popular or high-achieving, a student with low social self-concept is likely to acquiesce to the position argued by her/his opponent.

Alternative reactions

By discriminating between observable (overt) behaviour (keeping one's position during the discussion or giving it up) and cognitive (covert) processes (changing one's real belief or keeping it) we may identify four possible reactions in a discussion: (a) claim to stick to one's position and really believe in it; (b) claim to stick to one's position but lose faith in it; (c) overtly give up one's position but still believe in it; and (d) overtly give up one's position and really change the position. These four reactions will be referred to as (a) no overt change—no covert change (NOC-NCC), (b) no overt change—covert change (NOC-CC), (c) overt change—no covert change (OC-NCC), and (d) overt change—covert change (OC-CC). The four reactions are shown in Table 1.

TABLE 1

MODEL OF OVERT AND COVERT REACTIONS IN A DISCUSSION

		Overt reactions	
		No change/ keep position	Change/ give up position
Covert reactions	No change	NOC-NCC	OC-NCC
	Change	NOC-CC	OC-CC

Students with high academic achievement and high academic self-concept were predicted to react more by NOC-NCC and less by OC-CC and OC-NCC than students with low academic achievement and low self-concept. The latter reaction means overtly giving up one's position even if one continues to believe in it. We regard this reaction to be highly self-protective.

Reacting by NOC-CC, which means not giving up one's position during a discussion even if one changes one's mind, seems harder to predict. This reaction may be anticipated if some highly secure and competitive students enjoy a discussion by its own right. On the other hand such a self-presentation may also be expected by some self-protective students. The purpose of this strategy may not be to lose face.

Students with low self-concept were predicted to react more by OC-NCC and less by NOC-NCC than students with high social self-concept.

The present experiment was conducted to test these predictions in a situation where dyads of students discussed a problem with alternative solutions and each student was told which solution to argue for. There is normally a high correlation between ability and academic achievement. As a rule, high achievers will have better intellectual abilities than low achievers. In a spontaneous discussion high achievers might therefore be expected to be more able than low achievers to find the best and most persuasive arguments. This represented a methodological problem. To minimise the impact of ability to find good arguments each student was given four statements in favour of the position he/she was asked to argue for.

METHOD

Sample

The sample consisted of 80 students from ten 8th grade classes in three Norwegian city schools. All students in the ten classes were rated by the teachers and classified in three equally large groups as high achievers, medium achievers, and low achievers. Eight students from each class — two high and two low achievers of each sex — were then drawn for the experiment by a random stratified procedure to form four types of student dyads. The four types were (a) high achieving male — low achieving male (HM-LM), (b) high achieving female — low achieving female (HF-LF), (c) high achieving male — low achieving female (HM-LF), and (d) high achieving female — low achieving male (HF-LM). Thus, the experiment was conducted with 40 dyads of students — 10 dyads of each type.

Procedure

Before the study started academic self-concept, social self-concept and global self-esteem were measured by a questionnaire completed by all students in the ten classes. The questionnaire was administered by three trained research assistants. The students were instructed to work by themselves, and told that there were no right or wrong answers to the items. After careful instruction each item was read aloud by the assistant.

In each class the experiment was conducted by taking two pairs of students out of the classroom at a time. The dyads were placed with one experimenter in each of two separate rooms. When the students returned to the classroom, the two other dyads from the same class were taken out immediately to prevent the students from getting prior information about the experiment.

The experimental procedure required a text specifying a problem and stating alternative solutions for discussion. As part of the planning of the experiment three different texts focusing different problems were made. All of these texts stated a problem, gave two alternative solutions and provided four arguments for each solution. The texts were tested on an 8th grade class. The text used in the experiment was chosen for three reasons: (a) it seemed to be the text that was most interesting to the students; (b) it was the text coming closest to 50 per cent of the students preferring each of the two alternative solutions (55 per cent preferred alternative A while 45 per cent preferred alternative B), and (c) the students indicated that the text and the arguments were not difficult to understand.

The experiment started with the experimenter reading the following text twice:

> "A Norwegian community, which we might call X, is in a very difficult financial position. The expenses have to be cut down. To save money two suggestions have been made. The first alternative is to close down one of the departments at the local hospital and to cut down the financial support given to pensioners who pay high rents for their apartments. In other words: this alternative means reducing the community's grants to health and social purposes. It will not be necessary to fire any of the staff members, as they have kept the department running by working overtime. Thus, this solution means that the sick and the elderly people will suffer. More people will have to wait for treatment at the hospital and the pensioners will have a hard time.
>
> The other suggestion is to stop the financial support of a factory which is situated in the community. The business has given a deficit for years, and the factory has been dependent on financial support from the community. The people working at the factory pay taxes, but this tax income does not compensate for the support given by the community. By stopping the support to the factory the community will save the money it needs. As a consequence the factory will have to fire 100 workers. This means increased unemployment benefits, but still the community will save money.
>
> The two ways to save money for the community are: (1) to close a department at the hospital and reduce the support given to meet the heavy renting expenses for the elderly people. The sick and the elderly people will suffer. (2) to stop the financial support of the factory. Many workers will become unemployed, and the workers and their families will suffer. There are no other ways to solve the financial problems in the community."

The students were allowed to ask questions to make sure they had understood the text. When needed, part of the text was read again.

The students were then given the following instruction:

"As you will see, the community has to spend less money, and they have to decide how to do it. There are only two alternatives. You are now asked to argue for one of these. 'John' will argue for alternative A, and 'Mary' will argue for alternative B. I shall give each of you some arguments for the position you are asked to take."

The students were then given four written arguments for the alternative they were assigned to and they got 10 minutes to study the arguments. During this period the students were allowed to ask questions. The arguments were simple in order for all students to understand them, and none of the students seemed to have any problems in understanding and using them. The instruction then continued:

"You are now going to discuss what the community should do. Remember that there are only two possible solutions. No matter what alternative you think is the better, each of you must start arguing for the position you are given, using your own arguments and/or the written arguments you have received. However, you did not choose the alternative you are going to argue for yourself — it was given to you. It is therefore important to listen to the arguments for the other alternative. During the discussion you may be convinced that the alternative that you are asked to argue for is not the better. In that case you should give up the discussion and claim that the alternative solution is the better one. Since you have not chosen the position yourselves, giving up a position is no failure. The goal of this experiment is to find out what solution 8th grade students think is the better after having discussed the problem."

The discussion was stopped after 15 minutes if none of the students had given up their position by then. The experimenter noted whether any of the students gave up their position and the duration of the discussion. After having finished the discussion, the students completed a form indicating which alternative they really thought was the better (real position), independently of what they had claimed during the discussion. In order to enhance the validity of this measure of their real position after the discussion, the research assistants made the two students fill out the forms independently of each other and without any comments on what they really thought. The students also marked how certain they felt about what was the better solution.

Design

By a random procedure alternative A was given to five high achievers and to five low achievers in each of the four types of dyads, which constituted a total of eight categories of pairs. The categories and the number of pairs in each category is shown in Table 2.

TABLE 2
EXPERIMENTAL CATEGORIES AND NUMBER OF PAIRS IN EACH CATEGORY

Categories	N of Pairs
HM.A - LM.B	5
HM.B - LM.A	5
HF.A - LF.B	5
HF.B - LF.A	5
HM.A - LF.B	5
HM.B - LF.A	5
HF.A - LM.B	5
HF.B - LM.A	5

NOTE: H = high achiever, L = low achiever, M = male, F = female, A = asked to argue for alternative A, B = asked to argue for alternative B.

Instruments

General academic self-concept was defined as the students' general feeling of doing well in school and their satisfaction with their achievement. It was measured by a seven-item Academic Self-Esteem Scale (Skaalvik, 1986a, 1986b, in press) modelled after Harter's (1979) Perceived Competence Scale for Children. In neutral terms each item described two different groups of students, and the students were asked to indicate what group they were most like on a four-point scale. The rationale for this format was to make all answers look equally legitimate, thereby decreasing the social expectancy effect on the answers. The items were related to school achievement, but not to any particular subjects. The seven items are described in more detail elsewhere (Skaalvik, in press).

The reliability of the Academic Self-Esteem Scale has been demonstrated in several studies. Cronbach's alpha in four different studies has been shown to be between 0.72 and 0.78, while the test-retest reliability has been found to be 0.70 (Skaalvik, 1986b, in press). Cronbach's alpha in the present study was 0.83.

Global self-esteem was defined as an individual's general self-acceptance or his general positive or negative attitudes towards himself. Thus, high self-esteem implies that the individual sees himself as a person of worth, though he does not necessarily believe he is superior to others, while low self-esteem implies self-rejection, self-dissatisfaction, or self-contempt (Rosenberg and Simmons, 1975). Global self-esteem was measured by a context-free 13-item Self-acceptance Scale modelled after Harter's seven-item self-esteem scale (Harter, 1979) and has previously been used by Skaalvik (1983, 1986b, in press) and Skaalvik and Hagtvet (1990). The items were designed the same way as the Academic Self-Esteem Scale. The scale and measures of reliability showing test-retest correlation of 0.78 and Cronbach's alpha of 0.80 and 0.81 is presented elsewhere (Skaalvik, 1983, 1986b, in press; Skaalvik and Hagtvet, 1990). Cronbach's alpha in the present study was 0.75.

Social self-concept was defined as the students' feeling of being accepted by their classmates. It was measured by a four-item/four-point Social Self-Concept Scale measuring the students' feeling of being bullied by classmates, being lonely at school, having good friends in the class, and having difficulties getting contact with their classmates. Cronbach's alpha has previously been found to be 0.74 (Skaalvik, 1986a) and was 0.62 in the present study.

Real position was defined as the students' judgments of which alternative they "really" preferred, independent of what they had said in the discussion. Each student marked her/his real position after the discussion session by crossing either alternative A or B.

RESULTS

Overt reactions (claiming to change one's mind and stop the discussion) were significantly influenced by achievement in school and by academic self-concept (Tables 3 and 4). Low achieving students and students with low academic self-concept gave up the discussion and claimed to defer to the other person's position more often than did high achieving students and students with high self-concept. A regression analysis revealed that academic self-concept had no direct effect on overt reactions (beta = 0.02, $P>0.05$), while achievement in school had a direct and significant effect (beta = 0.317, $P = 0.05$). Sex of subject, global self-esteem and social self-concept had no significant impact on the students' overt reactions ($\chi^2 = 0.025$, $P>0.05$, $\chi^2 = 3.088$, $P>0.05$ and $\chi^2 = 0.705$, $P>9.05$ respectively).

None of the independent variables had any significant impact on the students' covert reactions (what they really believed to be the best solution after they had finished the discussion). There was a tendency at the 10 per cent level of significance that covert reactions were influenced by global self-esteem ($\chi^2 = 4.662$, df = 2, $P<0.10$) and by academic self-concept ($\chi^2 = 4.590$, df = 2, $P = 0.10$) in such a way that students with low self-concepts more easily than students with high self-concepts lost confidence in their positions. However, a regression analysis revealed that none of the independent variables had any significant direct effect even at the 10 per cent level (beta-values for achievement, academic self-concept, global self-esteem, and social self-concept were: 0.095, -0.097, -0.112, and 0.095 respectively).

TABLE 3

OVERT REACTIONS FOR HIGH AND LOW ACHIEVERS: RAW DATA

Overt reaction	Achievement	
	Low	High
NOC	20	33
OC	20	7
Total	40	40

NOTE: NOC = no overt change, OC = overt change.
$\chi^2 = 8.050$ df = 1 $P<0.01$

TABLE 4

OVERT REACTIONS FOR HIGH AND LOW ACADEMIC SELF-CONCEPT: RAW DATA

Overt reaction	Academic Self-concept		
	Low	Medium	High
NOC	15	16	21
OC	16	8	3
Total	31	24	24

NOTE: NOC = no overt change, OC = overt change.
$\chi^2 = 9.210$ df = 2 $P<0.01$

Overt and covert reactions were combined to four patterns of reactions as shown in Table 1. The impact of the independent variables on these patterns of combined reactions were tested by measures of chi-square. These tests revealed no significant impact of sex of subject ($\chi^2 = 4.296$, df = 3, $P>0.05$), academic self-concept ($\chi^2 = 11.222$, df = 6, $P>0.05$), global self-esteem ($\chi^2 = 6.888$, df = 6, $P>0.05$), or social self-concept ($\chi^2 = 1.613$, df = 2, $P>0.05$). However, there was a significant impact of achievement in school, as shown in Table 5. High achieving students reacted more than low achieving students by neither overtly giving in during the discussion nor covertly changing opinion (NOC-NCC), while low achieving students reacted more than high achieving students by both overtly giving in and really changing opinion (OC-CC). Low achievers also reacted more by overtly giving in during the discussion even when they did not change opinion (OC-NCC).

The mean discussion time for the 27 pairs where one of the students gave up the discussion and claimed to change position was 9.7 minutes. There was no difference between high and low achievers in the time they took part in the discussion before claiming to change position (MS = 2.593, F = 0.211, $P>0.05$).

After having finished the discussion the subjects were asked what position they really thought was the better. Then they were asked how certain they were that they had found the better solution. Among those subjects who believed that the position they had been asked to argue for was the better, there was no difference in the level of certainty as a function of achievement in school (MS = 0.84, F = 1.448, $P>0.05$). Among the students who really had changed the position they originally were asked to argue for, high achievers were more certain of their positions than low achievers. The difference was significant only at the 10 per cent level (MS = 2.100, F = 3.223, P = 0.08).

TABLE 5

OVERT AND COVERT REACTIONS FOR HIGH AND LOW ACHIEVERS: RAW DATA

Reactions	Achievement	
	Low	High
OC-NCC	6	2
NOC-NCC	13	24
OC-CC	14	5
NOC-CC	7	9
Total	40	40

NOTE: OC = overt change, NCC = no covert change,
NOC = no overt change, CC = covert change.
$\chi^2 = 9.783$ df = 3 $P<0.05$

DISCUSSION

We found no clear evidence that achievement in school had any impact on the students' cognitions related to the problem being discussed. On the other hand, we found that level of achievement strongly affected overt behaviour in the discussion. Academic self-concept had no direct impact on discussion behaviour or on the students' cognitions, and controlling for the effect of academic self-concept did not change the impact of achievement.

The lack of a direct effect of academic self-concept was surprising and calls for further verification. A possible post hoc interpretation is that the non-involvement strategy is not a result of low academic self-concept, but a protective strategy used by low achievers in situations that are threatening to their self-esteem. Another interpretation might be in terms of social roles. The role of low achieving students in academic situations may be a passive and non-involving one, seemingly accepting the judgments of others, more or less independently of their beliefs about the subject matter and about themselves. Future research should be designed to test these interpretations.

The results have relevance to the preparation of students for participation in democratic processes and to the functioning of such processes in a modern society. Even though consensus within a school on some important value questions is important for the education of the students, schools should take steps to ensure that students are confronted by different values and opinions. If not, they may lose the freshness of view that such differences can bring, and at worst, the students may be unequipped to enter the heterogeneity of the larger society (Coleman, 1987). It has further been suggested that the survival of a pluralistic democracy depends on a belief that mutual understanding can be achieved through genuine dialogue and critical encounter, and that it is important for a democracy to develop commitment of its citizens to such discourse (Bryk, 1988). This points to the importance of active participation in discussions in school. Our results suggest that low achieving students may play a negligible role in discussions. This means that their school experiences may not prepare them for democratic functioning. Whatever the school gives of information about democratic processes surely does not compensate for insufficient or maybe even negative experiences of low achieving students in discussion situations.

The present experiment was a first attempt to address the relationship between academic performance and behaviour in discussion situations. The results should be regarded as tentative and the study should be replicated with larger samples. Some limitations in the generalisation of the present data should be pointed out. We did not explicitly measure the students' understanding of the problem or the arguments they were given. Therefore, although the students made meaningful use of the arguments, low achievers may have had a poorer understanding of both the problem and the arguments than high achievers. The students' real position after the discussion was measured by self-report. It is not clear if and how the students' behaviour in the discussion may have affected their responses. This constitutes a validity problem that should be addressed in future research using this methodology.

Correspondence and requests for reprints should be sent to Einar M. Skaalvik or Olav Sletta, Department of Education, University of Trondheim, 7055 Dragvoll, Norway.

REFERENCES

Breakwell, G. (1986). *Coping with Threatened Identities*. London: Methuen.

Brookover, W. B., and Passalacqua, J. (1981). Comparison of aggregate self-concepts for populations with different reference groups. In Lynch, M. D., Norem-Hebeisen, A. A., and Gergen, K. J. (Eds.), *Self-Concept. Advances in Theory and Research*. Cambridge, Mass.: Ballinger.

Bryk, A. S. (1988). Musings on the moral life of schools. *Am. J. Educ.*, 96, 256-290.

Coleman, J. S. (1987). The relations between school and social structure. In Hallinan, M. T. (Ed.), *The Social Organization of Schools. New Conceptualizations of the Learning Process*. New York: Plenum Press.

Covington, M. V. (1984). The self-worth theory of achievement motivation: findings and implications. *Elem School J.*, 85, 5-20.

Harter, S. (1979). *Perceived Competence Scale for Children*. Denver: University of Denver.

Kaplan, H. B. (1980). *Deviant Behavior in Defence of Self*. New York: Academic Press.

Kelman, H. (1961). Processes of opinion change. *Pub Opin. Q*, 25, 57-78.

Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *J educ. Psychol.*, 79, 280-295.

Morris, R. (1977). Increasing participation through the use of normative interventions. Unpublished doctoral thesis, Stanford University.

Rosenberg, M. (1968). Psychological selectivity in self-esteem formation. In Gordon, C., and Gergen, K. J. (Eds.), *The Self in Social Interaction*. New York: Wiley.

Rosenberg, F., and Simmons, R. (1975). Sex differences in the self-concept in adolescence. *Sex Roles*, 1, 147-159.

Shavelson, R. J. and Bolus, R. (1982). Self-concept: the interplay of theory and methods. *J. educ. Psychol.*, 74, 3-17.

Skaalvik, E. M. (1982). Selvoppfatning og skoleerfaringer. (Self-concept and school experiences). Unpublished manuscript, Department of Education, University of Trondheim.

Skaalvik, E. M. (1983). Academic achievement, self-esteem and valuing of the school — some sex differences. *Br. J. educ. Psychol.*, 53, 299-306.

Skaalvik, E. M. (1986a) Selvoppfatning hos skoleelever. (Self-concept in school children). Unpublished manuscript, Department of Education, University of Trondheim.

Skaalvik, E. M. (1986b). Age trends in male and female self-esteem in Norwegian samples. *Scand. J. educ. Res.*, 30. 107-119.

Skaalvik, E. M. (in press). Gender differences in general academic self-esteem and in success expectations on defined academic problems. *J. educ. Psychol.*

Skaalvik, E. M., and Hagtvet, K. (1990). Academic achievement and self-concept: an analysis of causal predominance in a developmental perspective. *J. Person. soc. Psychol*, 58, 292-307.

Stulac, J. (1975). The self-fulfilling prophecy: modifying the effects of a unidimensional perception of academic competence in task-oriented groups. Unpublished doctoral dissertation, Stanford University.

(*Manuscript received* 11*th July*, 1989)

Br. J. educ. Psychol., 60, 142-152, 1990

SOME ANTECEDENTS OF ACADEMIC SELF-CONCEPT: A LONGITUDINAL STUDY

BY JAMES W. CHAPMAN AND ROBERT LAMBOURNE
(Massey University, New Zealand)
AND PHIL A. SILVA
(Dunedin Multidisciplinary Health and Development Research Unit,
University of Otago, New Zealand)

SUMMARY. The influence of cognitive, achievement, and family background variables on academic self-concept was examined for 435 students participating in a major longitudinal study in New Zealand. Family background variables included mother's marital status at the birth of the child, family socio-economic status at the time the child was born, and family environment when the child was 7 years and 9 years of age. These factors were not significantly related to academic self-concept at ages 9 and 11. In addition, the general emotional status of mothers when their child was 3 years and 5 years old was not significantly related to academic self-concept at ages 9 and 11. Mother's intelligence when the child was 3 years, and mother's expectations for school performance when the child started school at age 5 were not significantly related to academic self-concept. Similarly, cognitive and intellectual functioning at ages 3, 5, 7, and 9 years were not significantly related to academic self-concept at ages 7 and 9 years. Rather, academic self-concept at age 11 appeared to be influenced primarily by reading achievement and academic self-concept at age 9, and academic self-concept at age 9 was influenced primarily by reading achievement at age 7. It was concluded that academic self-concept is mainly a product of school achievement, and that any influence home background factors may have are likely to occur by means of their effect on school achievement.

INTRODUCTION

Interest in self-referent constructs has increased over the last two decades. This interest is particularly evident in areas of self-concept (Wylie, 1979; Burns, 1982; Byrne, 1984). Self-perceptions help define for individuals the nature of their relationships with other people, and the types of behaviours and tasks in which they will engage, (e.g., Abramson *et al.*, 1978; Weiner, 1979).

Achievement behaviours have received considerable attention in terms of self-referent constructs, particularly self-concept (Byrne, 1984). Although numerous studies have investigated the relationship between general self-concept and learning outcomes, fewer have focused on academic self-concept and school achievement (Chapman, 1988, 1989). Yet, the more closely self-concept is associated with specific situations, the closer is the relationship between self-concept and behaviour in the specific situation (Shavelson *et al.*, 1976). Thus, academic self-concept has a stronger relationship with school achievement than does general self-concept with achievement (Byrne, 1984; Chapman, 1988).

Having demonstrated the link between academic self-concept and achievement, attention has turned to influences that mediate that relationship. Song and Hattie (1984), for example, found that family structure and social status are related to each other and to family psychological characteristics. These family characteristics are in turn associated with self-concept, "particularly academic self-concept" (Song and Hattie, 1984, p. 1279).

If home background and developmental factors are important influences on the self-concept, questions arise as to when and how they affect self-concept. For example, do they have a direct influence? Or do family factors influence self-concept via some other variable such as intelligence or achievement? Another question that arises relates to when family influences show clear relationships with self-concept. Are these apparent when the child is young, or do they show up as the child progresses through school? Alternatively, insofar as school achievement and academic self-concept are "reciprocally interactive" (Bandura, 1977), it may well be that home/parental and developmental factors diminish in their influence on academic self-concept once school experience occurs. Actual achievement for example, may quickly override home and developmental influences on academic self-concept. In this regard, Song and Hattie (1984) found that family psychological characteristics, which are influenced mainly by social status, indirectly affect achievement through presentation of self. Presentation of self involves both physical self-concept and self-confidence, and in turn is associated with academic self-concept, which influences achievement.

Research reported to date has provided useful information on factors involved in the link between achievement and self-concept, especially academic self-concept. But most studies have involved concurrent data. There appear to be no studies that have focused on academic self-concept and achievement in terms of antecedent family, cognitive, and achievement variables. The "antecedent" studies that have been reported involved concurrent data and/or have focused on general self-concept or self-esteem (e.g., Rosenberg, 1965; Coopersmith, 1967; Sears, 1970). The Sears (1970) study for example, appears to be one of the few longitudinal studies of self-concept development, but its focus was on general self-concept. Similarly, Coopersmith's (1967) study of self-esteem covered a global range of factors, in which the antecedent elements related to concurrent background and developmental correlates.

The present study examined the associations between academic self-concept, reading performance (as a primary indicator of school achievement), and antecedent cognitive and family background measures over time. Data relating to these variables were available for a large sample of children participating in the longitudinal Dunedin (New Zealand) Multidisciplinary Health and Development Study.

Specifically, home background factors included socio-economic status (SES). This variable is commonly used to describe families, and correlates positively with intelligence and school achievement (Silva *et al.*, 1982; White, 1982). Song and Hattie (1984) found that social status was a significant influence on academic achievement through its effect on family psychological characteristics and self-concept.

Father-presence has also been considered an important influence on self-concept development. It is commonly assumed that "intact" homes are more favourable for children's psychological development than "broken" homes (Wylie, 1979). Bain *et al.* (1983) found that by Grade 3 father-present children showed more self-confidence in their academic ability, and were achieving at higher levels than father-absent children.

Family psychological factors have been studied in terms of self-concept. These include encouragement and expectations, along with educational interests and activities (Song and Hattie, 1984). Family interest in cultural, intellectual and political activities would seem to contribute to children's achievement-related perceptions and performance in school. Fotheringham and Creal (1980), for example, found that "general opportunities for learning", "the educational atmosphere of the home", and "parental expectations for the child" were significantly associated with cognitive development and academic achievement. Similarly, Share *et al.* (1983) found that SES and "home process" measures were significantly related to reading performance in children. These effects of family environment on school achievement are thought to occur through self-concept (Song and Hattie, 1984). Thus, families which value and foster an interest in academic activities probably value and expect competence in school. Such values and expectations may be associated with the development of positive school-related self-perceptions, which in turn influence achievement outcomes.

Finally, cognitive and achievement variables were also examined in this study. Research indicates that such school-related factors are among the strongest predictors of academic self-concept (e.g. Shavelson and Bolus, 1982; Byrne, 1984). However, academic self-concept as a function of antecedent cognitive/achievement variables does not appear to have been observed over an extended period of time.

In sum, this study examined longitudinal data in terms of academic self-concept, academic achievement, cognitive/intellectual characteristics, and family background variables. The aim was to ascertain the influence of the cognitive, achievement, and family factors on school ability-related perceptions.

METHOD

Sample

The sample consisted of 435 children who were a subset of children enrolled in the Dunedin Multidisciplinary Health and Development study. These children have been studied at birth, 3 years, 5 years, 7 years, 9 years and 11 years. A detailed sample description has been presented by McGee and Silva (1982). In summary, 1139 children born at Dunedin's one maternity hospital (Queen Mary) between 1 April 1972, and 31 March 1973, constitute the cohort for a multidisciplinary longitudinal study. Sample attrition is described by Williams and Silva (1985). Drop-out rates were relatively small up to and including the eleven-year assessment time (10 per cent). Those lost to the sample did not differ in socio-economic status from those retained in the sample. It is known that the sample is representative of Dunedin children, but slightly socioeconomically disadvantaged in comparison with New Zealand as a whole. Williams and Silva (1985) also note that the sample is underrepresentative of Maori and other Polynesian children (2 per cent) compared with New Zealand as a whole (10 per cent).

Missing data in the cohort are due to sample attrition at various stages of the study, lack of co-operation of children for certain tests, and incomplete results because some assessments were carried out by psychologists in other parts of the country who were not trained in the administration of all the tests used at the research unit. In addition, a printing error resulted in the omission of one page of the academic self-concept measure used at age 11. This omission led to some 520 children not being included in the present study.

Nonetheless, the subsample examined in this study is representative of the cohort still participating in the longitudinal research. For example, there were no significant differences between children assessed and those not assessed at age 11 in terms of academic self-concept, reading, and IQ scores at age 9.

Measures

Academic self-concept was assessed at ages 9 and 11 with the Perception of Ability Scale for Students — Short Form (PASS: Boersma and Chapman, in press). This scale was previously known as the Student's Perception of Ability Scale (SPAS: Boersma and Chapman, 1977, 1984). Thirty-five of the 70 items in the PASS comprise the Short Form. These items were selected on the basis of their subscale factor loadings. Half the number of items in each subscale, specifically those with the highest loadings, were included in the Short Form. The subscales are Perception of General Ability, Perception of Arithmetic Ability, General School Satisfaction, Perception of Reading and Spelling Ability, Perception of Penmanship and Neatness (each containing six items in the Short Form), and Confidence in Academic Ability (five items in the Short Form). Technical data for the PASS Short Form are reported by Chapman *et al.* (1983).

Parental and family characteristics were assessed with a number of measures. SES was determined at the time the children were born with the Elley and Irving (1972) index. This SES index was designed for use in New Zealand, and families are assigned to one of six levels on the basis of the father's occupation. In addition, mother's marital status at the time of their child's birth was also noted.

The intellectual and cultural orientation of the family was assessed by means of the relevant scales from the Family Environment Scale (FES: Moos, 1974). This scale deals with family interest in cultural, political, and intellectual activities. The questionnaire was completed by mothers when the children were aged 7, and again when they were aged 9 years (Parnicky *et al.*, 1985).

Another indicator of home are characteristics of the parents. Some characteristics of the mothers were assessed with the Malaise Inventory (Rutter *et al.*, 1970). This questionnaire was designed to sample a range of symptoms commonly associated with emotional disturbance in adults. Silva *et al.* (1982) report the results of the use of this inventory with the mothers of the Dunedin sample. They show that Malaise Inventory scores were significantly predicted by low SES, being a young mother at first pregnancy, separation from her partner, and having sought marriage guidance counselling. The inventory is a useful indicator of general emotional well-being in the mother.

Maternal intelligence was assessed when the children were 3 years of age by use of the SRA Verbal Test (Thurstone and Thurstone, 1973). Results from the use of this test with the Dunedin sample are described by Silva (1978).

When the children were assessed at age 5, their mothers were asked how they expected their child to perform academically at school. Responses were recorded in terms of three categories: below average, average, and above average.

Cognitive and achievement levels were assessed with different tests at different ages. Three measures of intelligence were used. At age 3, the Peabody Picture Vocabulary Test (Dunn, 1965) was administered, while at age 5, the Stanford-Binet Intelligence Scale (Terman and Merrill, 1960) was used. At ages 7 and 9 the Wechsler Intelligence Scale for Children-Revised (WISC-R: Wechsler, 1974) was administered. Because of time constraints, however, only eight subscales of the WISC-R were used. Comprehension was deleted from the Verbal Scale. IQ scores were prorated using the method described in the test manual.

At ages 3 and 5, Verbal Comprehension and Expression were assessed by the Reynell Developmental Language Scales (Reynell, 1969). These scales assess verbal comprehension, and appraise the child's understanding of simple instructions. In addition each child's spoken language in a number of situations is assessed in terms of language structure, vocabulary and language content. The Goodenough-Harris Draw-a-Man Test (Harris, 1965) was administered when the children were 5 years old in order to provide an estimate of non-verbal intelligence. Speech articulation was also assessed at age 5, by means of the Dunedin Articulation Screening Scale (Silva, 1980).

A factor analysis using the principal components method with iterations (SPSS Inc., 1986) was carried out on the three measures of IQ, verbal comprehension, and verbal expression at age 3. The factor loadings were used to make up a composite cognitive ability score. Similarly, a composite cognitive ability score at age 5 was derived from factor loadings arising from a principal components analysis of IQ (Stanford-Binet), verbal comprehension and expression, draw-a-man, and speech articulation tests. Results of these factor analyses are more fully reported by Williams and Silva (1985).

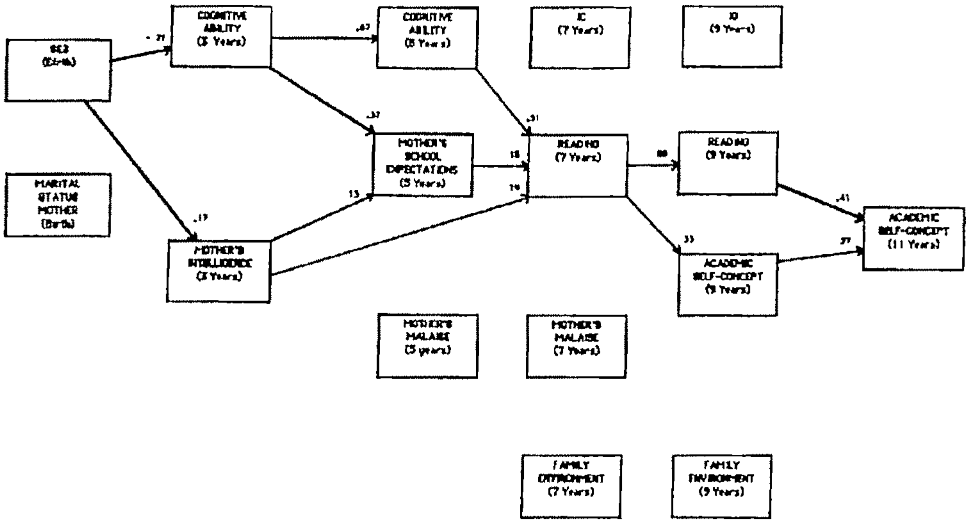
Finally, an indicator of *school achievement* at ages 7 and 9 was obtained using the Burt Word Reading Test (Scottish Council for Research in Education, 1976). This test provides an estimate of word recognition skills and thus gives an indication of superiority or backwardness in reading, a key school subject. It is suitable for use with children from age 5 to 13 years. Data from a New Zealand standardisation study (Gilmore *et al.*, 1981), indicate that the Burt Reading Test correlates highly with measures of reading vocabulary and reading comprehension. Scores have been found to correlate in the range of 0.72 to 0.82 with a broader measure, the Test of Scholastic Abilities (Reid *et al.*, 1981). This test was designed as a broadly based assessment of scholastic abilities, especially those verbal and numerical reasoning abilities considered to be required for success in academic subjects in New Zealand schools.

Statistical analyses

In order to explore the contribution of the variables to academic self-concept, a path model was postulated and tested using path analysis. The development of the model involves the statement of hypotheses relating to the influence of variables within the constraints of the directionality set down in the model. The validity of the model is tested by the analyses. These analyses produce a series of path coefficients. The magnitude of the path coefficients suggests the strength of relationships between variables. The path coefficients used to describe the model were standardised beta weights derived from a series of multiple regressions. The description and analysis of the model are in line with strategies adopted by various other researchers (e.g. Wolfe, 1980; Keith, 1982; Marsh and Parker, 1984). The path model explored in this study is set out in Figure 1.

FIGURE 1

PATH DIAGRAM SHOWING STRUCTURE OF RELATIONSHIPS BETWEEN ACADEMIC SELF-CONCEPT AND BACKGROUND FAMILY AND COGNITIVE/ACHIEVEMENT VARIABLES.


SES (Birth)
COGNITIVE ABILITY (3 Years)
COGNITIVE ABILITY (5 Years)
IQ (7 Years)
IQ (9 Years)
MARITAL STATUS MOTHER (Birth)
MOTHER'S SCHOOL EXPECTATIONS (5 Years)
READING (7 Years)
READING (9 Years)
MOTHER'S INTELLIGENCE (3 Years)
ACADEMIC SELF-CONCEPT (9 Years)
ACADEMIC SELF-CONCEPT (11 Years)
MOTHER'S MALAISE (5 years)
MOTHER'S MALAISE (7 Years)
FAMILY ENVIRONMENT (7 Years)
FAMILY ENVIRONMENT (9 Years)
.21
.67
.37
.31
.17
13
18
14
88
.41
33
27

Beta weights are shown on each path. Paths have been drawn only where the beta weights are both greater than 0.10 and statistically significant at the 0.05 level.

RESULTS

Correlations

Table 1 sets out the matrix of correlations among the variables which feature in the analysis. Of particular interest in this study were the correlations for PASS scores at ages 9 and 11 with the other variables. It is clear from the correlations presented in Table 1 that the relationship between home background variables and academic self-concept was negligible.

Only mother's school expectations at age 5, mother's SRA score at age 3, and SES at birth showed statistically significant relationships with PASS scores at ages 9 and 11. Perhaps of greater importance were the zero correlations between the intellectual and cultural orientation of the family. These variables were assessed at ages 7 and 9, yet were not associated with academic self-concept.

TABLE 1

PRODUCT-MOMENT CORRELATIONS AMONG THE VARIABLES IN THE STUDY (N = 435)

Measures	Age of Child	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Family Background																
1. Mother's marital status	birth															
2. SES	birth	-0.059														
3. Mother's IQ	3	-0.026	-0.167													
4. Mother's malaise	5	0.095	0.217*	-0.253*												
5. Mother's school expectations	5	0.015	-0.225*	0.199*	-0.107											
6. Mother's malaise	7	0.107	0.140	-0.191	0.623*	-0.096										
7. FES	7	-0.047	0.022	0.069	0.050	0.014	0.081									
8. FES	9	0.075	0.043	-0.024	0.070	-0.042	0.060	0.525*								
Cognitive/Achievement																
9. Cognitive Ability	3	0.059	-0.210*	0.216*	-0.156	0.369*	-0.144	-0.001	-0.082							
10. Cognitive Ability	5	-0.014	-0.225*	0.239*	-0.144	0.448*	-0.113	0.013	-0.004	0.670*						
11. WISC-R IQ	7	0.003	-0.185	0.326*	-0.210*	0.452*	-0.156	0.041	-0.012	0.526*	0.613*					
12. Burt Reading	7	0.008	-0.173	0.302*	-0.167	0.387*	-0.127	-0.012	-0.020	0.394*	0.508*	0.509*				
13. WISC-R IQ	9	-0.008	-0.201*	0.340*	-0.209*	0.427*	-0.176	0.014	-0.017	0.558*	0.655*	0.779*	0.548*			
14. Burt Reading	9	0.010	-0.196*	0.288*	-0.161	0.370*	-0.111	0.012	0.034	0.346*	0.517*	0.510*	0.876*	0.555*		
15. PASS	9	0.006	-0.008	0.102	-0.051	0.128	-0.003	-0.008	0.002	0.153	0.240*	0.179	0.340*	0.215*	0.321*	
16. PASS	11	0.041	-0.133	0.144	-0.081	0.163	-0.060	-0.022	-0.011	0.182	0.257*	0.209*	0.397*	0.282*	0.407*	0.376*

* P < 0.05

SES = Socio-economic status of family
FES = Family Environment Scale
PASS = Perception of Ability Scale for Students

TABLE 2

PREDICTIVE CORRELATIONS[1] BETWEEN ACADEMIC SELF-CONCEPT SCORES AND FAMILY BACKGROUND AND COGNITIVE/ACHIEVEMENT MEASURES (N = 435)

	Age	PASS (11 yrs)	PASS (9 yrs)
Family Background			
Mother's marital status	birth	0.036	0.003
SES	birth	-0.071	0.053
Mother's intelligence	3	-0.006	0.001
Mother's malaise	5	-0.016	0.008
Mother's school expectations	5	0.011	0.001
Mother's malaise	7	-0.024	0.042
Family environment	7	-0.024	-0.003
Family environment	9	-0.022	
Cognitive/Achievement			
Cognitive ability	3	0.033	0.028
Cognitive ability	5	0.036	0.093
WISC-R IQ	7	-0.002	0.008
Burt reading	7	0.107	0.334
WISC-R IQ	9	0.068	
Burt reading	9	0.407	
Academic self-concept	9	0.273	
Multiple R		0.482	0 334

[1] Beta coefficients

In contrast to the above results, all cognitive and achievement variables were significantly correlated with PASS scores at each age. Of these, Burt Reading scores showed the strongest association with academic self-concept. Cognitive ability and IQ scores were less predictive of PASS scores than achievement, a finding that is consistent with other studies (Boersma and Chapman, in press).

The stability coefficient for PASS scores over the two-year period from age 9 to age 11 was 0.38. This is somewhat lower than that reported in other studies (Boersma and Chapman, in press). Chapman and Wilkinson (1988), for example, have reported correlations for PASS scores over a three- to four-year period that were around 0.50, and as high as 0.71. The Short Form of the PASS used in this study may be less stable over the time than the regular form.

Path analysis

A path model was developed in order to examine the possible causal ordering of variables over the time period of the study. This model is shown in Figure 1, and includes all variables in this study from birth through the two-year testing phases to age 11. The model is fully recursive and postulates a direction of causality from left to right. No causal ordering was hypothesised for variables that were temporally adjacent.

The results of the path analysis are summarised in Figure 1. The assumed causal effects are represented by standardised regression coefficients (Pedhazur, 1982). All variables tested in the full model have been included in the figure; however, non-significant paths, and those paths having standardised coefficients that were less than 0.10, have been excluded from Figure 1. The decision to delete paths whose coefficients were less than 0.10 follows Kerlinger and Pedhazur's (1973) comments regarding the meaningfulness of paths that may be small, yet statistically significant because of a large sample size.

The results show that the model had moderate precision in terms of explaining the variability of measures of academic self-concept in children. The multiple correlations indicate that the variables in the model accounted for 23.24 per cent of the variance of academic self-concept at age 11, and 11.14 per cent of the variance of academic self-concept at age 9. Despite the moderate precision of results, the findings are informative.

(1) Maternal socio-economic status and marital status at the birth of the child were not significantly correlated. However, maternal socio-economic status was significantly related to measures of the child's cognitive ability at age 3, and to mother's measured intellectual level when the child was 3 years.

(2) The child's cognitive ability at age 3 was significantly correlated with mother's intellectual level at the same time. Also, the measured cognitive ability of the child at age 3 correlated significantly with cognitive ability at age 5, and to the mother's expectations regarding school performance at age 5.

(3) The mother's level of intelligence as assessed when the child was 3, correlated significantly with her school expectations for the child at age 5, and with the child's reading performance at age 7.

(4) The child's cognitive ability at age 5 was correlated with mother's school expectations at the same age, and correlated significantly with reading performance at age 7.

(5) Mother's school expectations for her child at age 5 were significantly associated with reading performance at age 7.

(6) Reading performance at age 7 correlated significantly with reading performance at age 9, and with academic self-concept at age 9.

(7) Reading performance at age 9 was significantly associated with academic self-concept at age 11, but the partial correlation with academic self-concept at age 9 was not significant ($r = 0.045$).

(8) Academic self-concept at age 9 correlated significantly with academic self-concept at age 11.

It is clear from Figure 1 that no family background or cognitive/intellectual variable had a direct effect on academic self-concept. Indeed, Family Environment Scale scores showed zero correlations with academic self-concept and achievement (see Table 1). In addition, Mothers' Malaise was not significantly associated with either academic self-concept or achievement.

Clearly, academic self-concept at age 11 was directly affected by reading ability and academic self-concept at age 9. The combined effect of these two nine-year variables was sizeable ($R = 0.48$; $F(2,431) = 65.25$, $P<0.001$). For both reading and academic self-concept at age 9, reading at age 7 was the only significant predictor (see Figure 1). Family variables and IQ had no direct significant effects on reading or academic self-concept. Reading performance at age 7 was clearly important in terms of its direct effect on academic self-concept at age 9, and its indirect effect on academic self-concept at age 11 through academic self-concept at age 9 ($0.09 = 0.33 \times 0.27$) and through reading at age 9 ($0.36 = 0.88 \times 0.41$). The total causal effect of reading at age 7 on academic self-concept at age 11 was 0.45 ($0.09 + 0.36$).

DISCUSSION

Numerous writers have claimed that research on the relation between self-concept and achievement should include environmental variables such as home background and parental influence, as either causal or moderating variables (e.g., Marjoribanks, 1979; Shavelson and Bolus, 1982). Although some studies have found the inclusion of such variables fruitful, the longitudinal data in this study do not support the belief that family background variables have any major, long-term effect on academic self-concept. Further, family interest in cultural, political, and intellectual activities were not associated with either academic self-concept or reading achievement. Academic self-concept was affected primarily by past achievement and past perceptions of ability.

These findings are at variance with those reported by Song and Hattie (1984). They found that the family psychological characteristics (e.g., parental education, educational activities in the home) of Korean high school students were causally related to self-concept variables, which in turn affected achievement. Their data, however, were collected concurrently. In the present study, home background variables appeared to have no direct or major indirect long lasting effects on achievement or academic self-concept.

If home environment does set the stage for self-concept development (e.g., Rosenberg, 1965; Coopersmith, 1967; Sears, 1970; Burns, 1982), then these effects may be more subtle than indicated by the measures in this study. Meanwhile, our results strongly support the hypothesis that academic self-concept is primarily the result of different levels of school achievement. This result is consistent with studies of the relationship between achievement and academic self-concept reported by Byrne (1984) and Hansford and Hattie (1982). Home environment effects on academic self-concept more likely occur through their influence on achievement during the very early stages of schooling. Once schooling starts, however, initially high school-related expectations and perceptions of ability begin to reflect actual achievement (Harter, 1983; Stipek, 1984).

The experience of schooling, involving comparison with others and feedback about academic performance, is likely to be the main factor in the development of an academic self-concept. From that time, the interplay between academic self-concept and achievement is probably reciprocal (Byrne, 1986), with academic self-concept being both a "cause" and "effect" of achievement. Stanovich (1986) refers to this reciprocal effect as the "Matthew effect", in which competent achievers do better over time, and poor achievers become more skill-deficient over time. Success or failure in schoolwork thus evolves into generalised "behavioural/cognitive/motivational spin-offs" (p. 389). For less skilled readers (for example), slower progress in reading has more generalised effects on processes that underlie a broader range of tasks and skills than just reading (Stanovich, 1986). Indeed, Butkowsky and Willows (1980) concluded that children with reading difficulties probably display an eroding motivation in achievement situations that increases the likelihood of failure in the future. The impact of such failure appears to occur within the first few years of school and appears to be consistent through to the end of high school (Chapman, 1988). The effect of low academic self-concept, and the frequently associated feelings of learned helplessness, may result in children giving up easily in the face of difficulty and never persisting long enough to discover that successes may be possible (Butkowsky and Willows, 1980).

In essence then, academic self-concept would seem to be closely linked to actual achievement outcomes in school. Any ongoing effects that home background may have on the reciprocal interplay between achievement and academic concept probably relate to the nature of achievement-related feedback and encouragement in the home. Future research on the nature of such interactions would seem useful in assisting the understanding of home/family effects on ability perceptions and achievement outcomes.

ACKNOWLEDGMENTS.—The Dunedin Multidisciplinary Child Development Study is supported by the Medical Research Council of New Zealand. Many of the data are gathered by voluntary workers from the Dunedin community. The authors are indebted to the many people whose contributions make this study possible.

Correspondence and requests for reprints should be addressed to Dr James Chapman, Education Department, Massey University, Palmerston North, New Zealand.

REFERENCES

Abramson, L. Y., Seligman, M. E. P., and Teasdale, J. (1978). Learned helplessness in humans: critique and reformulation. *J. abnorm. Psychol.*, 87, 49-74.

Bain, H. C., Boersma, F. J., and Chapman, J. W. (1983). Academic achievement and locus of control in father-absent elementary school children. *School Psychol. Int.*, 4, 69-78.

Bandura, A. (1977). Self-efficacy: toward a unifying view of human behavior. *Psychol. Rev.*, 84, 191-215.

Boersma, F. J., and Chapman, J. W. (1977). *The Student's Perception of Ability Scale.* Edmonton, Alberta: University of Alberta.

Boersma, F. J., and Chapman, J. W. (1984). *Manual for the Student's Perception of Ability Scale*, (rev. ed.). Edmonton, Alberta: University of Alberta.

Boersma, F. J., and Chapman, J. W. (in press). *Manual for the Perception of Ability Scale for Students.* Los Angeles: Western Psychological Services.

Burns, R. B. (1982). *Self-concept Development and Education.* London: Holt, Rinehart and Winston.

Butkowsky, I. S., and Willows, D. (1980). Cognitive-motivational characteristics of children varying in reading ability: evidence for learned helplessness in poor readers. *J. educ. Psychol.*, 72, 408-422.

Byrne, B. M. (1984). The general/academic network: a review of construct validation research. *Rev. educ. Res.*, 54, 427-456.

Byrne, B. M. (1986). Self-concept/academic achievement relations: an investigation of dimensionality, stability, and causality. *Can. J. behav. Sci.*, 18, 173-186.

Chapman, J. W. (1988). Learning disabled children's self-concepts. *Rev. educ. Res.*, 58, 347-371.

Chapman, J. W. (1989). The Perception of Ability Scale for Students: results from accumulated research. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, in March.

Chapman, J. W., and Wilkinson, L. (1988). *Self-Concept, Identity Status and Emotional Status of Form 5 Learning Disabled Students.* (New Zealand Department of Education Research Contract No. 47/17/151.) Palmerston North, N.Z.: Massey University.

Chapman, J. W., Silva, P., and Boersma, F. J. (1983). Student's Perception of Ability Scale: development of a Short Form. *Percept. mot. Skills*, 57, 799-802.

Coopersmith, S. (1967). *The Antecedents of Self-Esteem.* San Francisco: Freeman.

Dunn, L. (1965). *The Peabody Picture Vocabulary Test.* Minneapolis: American Guidance Service.

Elley, W. B., and Irving, J. C. (1972). A socio-economic index for New Zealand based on levels of education and income from the 1966 census. *N.Z. J. educ. Stud.*, 7, 153-167.

Fotheringham, J. B., and Creal, D. (1980). Family socio-economic and educational-emotional characteristics as predictors of school achievement. *J. educ. Res.*, 73, 311-317.

Gilmore, A., Croft, C., and Reid, N. (1981). *Burt Word Reading Test: New Zealand Revision Teachers' Manual.* Wellington: New Zealand Council for Educational Research.

Hansford, B. C., and Hattie, J. A. (1982). Relationship between self and achievement/performance measures. *Rev. educ. Res.*, 52, 123-142.

Harris, D. B. (1965). *Children's Drawings as Measures of Intellectual Maturity.* New York: Harcourt Brace.

Harter, S. (1983). Developmental perspectives on the self-system. In Mussen, P. (Ed.), *Carmichael's Manual of Child Psychology.* (Volume on Social Development, M. Hetherington, Ed.), New York: John Wiley and Sons.

Keith, T. Z. (1982). Time spent on homework and high school grades: a large sample path analysis. *J. educ Psychol.*, 74, 248-253.

Kerlinger, F. N., and Pedhazur, E. J. (1973). *Multiple Regression in Behavioral Research.* New York: Holt, Rinehart and Winston.

Marjoribanks, K. (1979). *Families and their Learning Environment.* London: Routledge and Kegan Paul.

Marsh, H. W., and Parker, J. W. (1984). Determinants of student self-concept: is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *J. Person. soc. Psychol.*, 47, 213-231.

McGee, R. O., and Silva, P. A. (1982). *A Thousand New Zealand Children: Their Health and Development from Birth to Seven.* Auckland: Medical Research Council of New Zealand. Special Report Series No. 8.

Moos, R. (1974). *Combined Preliminary Manual for Family, Work and Group Environmental Scales.* Palo Alto, Calif.: Consulting Psychologists Press.

Parnicky, J. J., Williams, S., and Silva, P. A. (1985). Family Environment Scale: a Dunedin (New Zealand) pilot study. *Aust. Psychol.*, 20, 195-204.

Pedhazur, E. J. (1982). *Multiple Regression in Behavioral Research* (2nd ed.), New York: Holt, Rinehart and Winston.

Reid, N. A., Jackson, P., Gilmore, A., and Croft, C. (1981). *Test of Scholastic Abilities.* Wellington: New Zealand Council for Educational Research.

Reynell, J. (1969). *Reynell Developmental language Scales.* Slough: National Foundation for Educational Research.

Rosenberg, M. (1965). *Society and the Adolescent Self-Image.* Princeton, N.J.: Princeton University Press.

Rutter, M., Tizard, J., and Whitmore, K. (1970). *Education, Health and Behaviour.* London: Longmans.

Scottish Council for Research in Education (1976). *The Burt Reading Test: 1974 Revision.* London: Hodder and Stoughton.

Sears, R. R. (1970). Relation of early socialization experiences to self-concepts and gender role in middle childhood. *Child Dev.*, 41, 267-289.

SHARE, D. L., JORM, A. F., MACLEAN, R., MATTHEWS, R., AND WATERMAN, B. (1983). Early reading achievement, oral language ability, and a child's home background. *Aust. Psychol.*, 18, 75-87.

SHAVELSON, R. J., AND BOLUS, R. (1982). Self-concept: the interplay of theory and methods. *J educ. Psychol.*, 74, 3-17.

SHAVELSON, R. J., HUBNER, J. J., AND STANTON, J. C. (1976). Self-concept: validation of construct interpretations. *Rev. educ. Res.*, 46, 407-441.

SILVA, P. A. (1978). SRA verbal test scores from 1011 women. *N.Z. Psychol.*, 7, 47-48.

SILVA, P. A. (1980). *The Dunedin Articulation Screening Scale*. Dunedin: Otago Speech Therapy Association.

SILVA, P. A., MCGEE, R. O., THOMAS, J., AND WILLIAMS, S. M. (1982). A descriptive study of social-economic status and child development in Dunedin five-year-olds: a report from the Dunedin Multidisciplinary Child Development Study. *N.Z. J. educ. Stud.*, 19, 21-32.

SONG, I-S., AND HATTIE, J. (1984). Home environment, self-concept, and academic achievement: a causal modeling approach. *J. educ. Psychol.*, 76, 1269-1281.

SPSS Inc. (1986). *SPSSx User's Guide.* (2nd ed.) New York: McGraw Hill.

STANOVICH, K. E. (1986). Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. *Read. Res. Q.*, 21, 360-406.

STIPEK, D. J. (1984). The development of achievement motivation. In AMES, R. AND AMES, C. (Eds.), *Research on Motivation in Education, Vol. 1: Student Motivation.* Orlando, FL: Academic Press.

TERMAN, L. M., AND MERRILL, M. R. (1960). *Stanford-Binet Intelligence Scale.* Boston: Houghton Mifflin.

THURSTONE, T. G., AND THURSTONE, L. L. (1973). *The SRA Verbal Form*, Chicago: Science Research Associates.

WECHSLER, D. (1974). *The Wechsler Intelligence Scale for Children-Revised.* New York: The Psychological Corporation.

WEINER, B. (1979). A theory of motivation for some classroom experiences. *J. educ. Psychol.*, 71, 3-25.

WHITE, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychol. Bull*, 91, 461-481.

WILLIAMS, S. M., AND SILVA, P. A. (1985). Some factors associated with reading ability: a longitudinal study. *Educ Res.*, 27, 159-168.

WOLFE, L. M. (1980). Strategies of path analysis. *Am. educ. Res J.*, 17. 183-209.

WYLIE, R. C. (1974). *The Self-Concept*, (rev. ed.). Lincoln: University of Nebraska Press.

WYLIE, R. C. (1979). *The Self-Concept, Vol. 2.* Lincoln: University of Nebraska Press.

(*Manuscript received 17th October*, 1988)

Br. J. educ. Psychol., 60, 153-160, 1990

HEADINGS WITHIN MULTIPLE-CHOICE TESTS AS FACILITATORS OF TEST PERFORMANCE

By Michael A. R. Townsend, Dennis W. Moore, Bryan F. Tuck, and Keri M. Wilton
(University of Auckland, New Zealand)

Summary. Undergraduate students (N = 287) taking a multiple-choice test were either provided with headings during the test or completed the test without headings. Students were also tested for field-independence and later completed a questionnaire concerning their beliefs about the effects of headings in tests. Headings did not facilitate test performance, nor was there any association with field-independence. However, the majority of students believed that headings would increase their test score as a function of activation of prior knowledge. Taken together with previous research, the findings suggest that the facilitative effects of headings occur at encoding rather than retrieval.

INTRODUCTION

A number of recent studies have indicated that student learning from prose text can be enhanced if headings are inserted ahead of sections of the text (Dee-Lucas and Di Vesta, 1980; Hartley and Trueman, 1983, 1985). In particular, several studies have shown that headings in prose text can have a facilitative effect on memory for text information as measured by multiple-choice tests given after learning (Brooks *et al.*, 1983; Wilhite, 1986, 1988). Headings appear to provide an organisational framework in which relevant prior knowledge is activated to facilitate encoding of information (Wilhite, 1989). However, little is known about the effects of headings at the time of retrieval, that is, within multiple-choice tests themselves.

Although there is no direct empirical evidence regarding the effects of headings for sections within multiple-choice tests, it seems likely that headings would facilitate test performance by activating prior knowledge schemata associated with each section of a test (Anderson *et al.*, 1978), and by assisting students in appropriate schema shifting between sections during the test (Townsend, 1983). Moreover, it also seems likely that these retrieval functions would be more enhanced if the headings in the test were also those used in the text. Such effects are not discussed in the literature on test construction, although Thorndike (1971) believed it desirable to identify groups of items by use of lines or asterisks, or by a "note" at the head of the group (he gives the following "note" as an example, "Items 6-9 apply to this diagram"). Furthermore, Wesman (1971) argued that grouping of similar items has been "found to enhance the quality of . . . performance" (p. 77), a claim given some limited empirical support by Flaugher *et al.* (1968) using random and grouped presentation of items of the Scholastic Aptitude Test. The current study sought to determine the effect of headings for groups of items on test performance.

The effect of test headings on performance may be influenced by cognitive style, particularly psychological differentiation. For example, field-independent students demonstrate an analytic ability to isolate the discrete elements of a stimulus and to impose organisation where none may be apparent, whereas field-dependent students have a more global style of processing that leaves them bound by the inherent organisation of the stimulus (Witkin *et al.*, 1977; Spiro and Tirre, 1980). A number of researchers have shown that although field-dependent students may have lower performance than their field-independent peers, their performance can be improved when they are provided with instructional support which delineates the structure of the learning material (Frank, 1984; Kardash *et al.*, 1988; Kiewra and Frank, 1988). Field-dependent students are most likely to benefit from headings in so far as headings provide an explicit organisational structure for the test. The current study sought to examine test achievement as a function of the interaction of headings and student psychological differentiation.

Although achievement is the major focus of most research on instructional strategies, there is a need for researchers to be cognisant of the relationship between the actual benefits of a strategy and the benefits as perceived by students. Cognisance of this relationship may affect the advice given to students taking multiple-choice tests; for example, on whether students should switch answers or persist with their first choice when uncertain (Skinner, 1983). But knowledge of perceived benefits may also influence the teacher's selection of a strategy. For example, Dodd and Leal's (1988) recommendation of an answer-justification strategy in multiple-choice tests (where a student may submit a short essay in lieu of completing an item about which there is uncertainty) rests on the finding that it is very popular among students. With regard to the current study, little is known about whether students believe that headings affect their test score, or how they might affect their test score, or whose test scores they might affect. Anecdotal evidence gathered by the first author following the inclusion of headings in a test given in a previous year suggested that students were highly positive about the effects of headings, but divided, for example, about whether headings were only beneficial if the material had been studied carefully. Based on these informal discussions, the current study investigated three questions regarding student beliefs about the effect of headings: (1) Are any facilitative effects of headings due more to assistance in remembering and understanding (a cognitive explanation) or to a reduction in test anxiety (an affective explanation)? (2) Are headings more likely to benefit well prepared students or poorly prepared students? (3) Should headings be used in multiple-choice tests?

In summary, the current study was designed to determine whether the inclusion of headings in a multiple-choice test would enhance the performance of students, particularly those with a field-dependent cognitive style, and whether any obtained effects are consistent with students' beliefs. These issues are important when considered against the frequent use of multiple-choice tests in assessment in educational settings.

METHOD

Sample

The sample comprised 287 undergraduate students enrolled in a second year course (25 weeks) in educational psychology at the University of Auckland. There were 237 females and 50 males.

Materials

Achievement was measured by a 41-item multiple-choice test based on the material for the first quarter of the course. (A 42nd item was deleted because it had been printed with no correct solution.) The items covered nine chapters from the first four sections of Gage and Berliner's *Educational Psychology* (1988), the course textbook. Some 37 of the test items were selected from the instructor's manual for the textbook, while the remaining four test items were constructed by the first author. Two forms of the achievement test were constructed. One form contained just the 41 items, while the other contained the same items, in the same order, but was divided into five sections of approximately equal numbers of items. Each section of the latter form began with a heading relating to the topics in the section. The headings were composed of the relevant chapter titles from the text book, each typed one per line in bold capitals. The five sections of the test referred to chapters 1, 2 and 3, chapters 4 and 5, chapter 8, chapters 12 and 13, and chapter 14, respectively. This division of the material followed the organisation presented to students in the course outline and followed in the lectures given by the first author. The coefficient alpha of internal consistency of the test was 0.74, based on the responses of the 287 students in the study.

Field-independence was measured by the Hidden Figures Test (HFT) (French *et al.*, 1969). The 32-item HFT is a moderately difficult test requiring the ability to recognise simple geometric forms embedded within complex stimuli. Kardash *et al.* (1988) found that undergraduate students in educational psychology needed a score of 12 or better to be in the upper one-third of scores (field-independent), and a score of 8 or less to be in the lower one-third of scores (field-dependent).

Students also completed a seven-item questionnaire concerning their beliefs about the effects of headings in multiple-choice tests. The items were designed to reflect students' beliefs about the effect of headings on (1) their test score, (2) their understanding of the items, (3) their remembering of information relevant to the items, (4) their test anxiety, (5) the scores of students who had studied adequately for the test, and (6) the scores of students who had not studied adequately for the test. The final item (7) asked whether headings should be used in multiple-choice tests. Students were asked to select one of three statements for each item, reflecting a positive effect (scored 1), no effect (scored 2), and a negative effect (scored 3). For example, Item 1 read:

> Having topic headings for the items in each section of a multiple-choice test would probably:
>
> (a) increase my test score
>
> (b) make no difference to my test score
>
> (c) decrease my test score.

At the conclusion of the seven items, students were given an open-ended opportunity to write any additional statements about the effects of headings in tests.

Procedure

The achievement test (worth 20 per cent of the final grade) was administered in a single one-hour supervised examination at the end of the first quarter of the course. The two forms of the multiple-choice test were randomly distributed to the 287 students present in the lecture hall. Scores on the Hidden Figures Test (N = 233) were obtained some three weeks before the achievement test as part of a small-group laboratory session exercise in the third week of term. The top 81 scores (HFT > 14) were classified as high in field-independence, the middle 81 scores (HFT = 8-14) as middle in field-independence, and the 71 low scores (HFT < 8) as low in field-independence. Finally, students' beliefs about the effects of test headings (N = 204) were assessed at small-group laboratory sessions the week following the examination; at this time they were unaware that there had been two forms of the achievement test. Testing in laboratory sessions was carried out by tutors responsible for teaching research design and data analysis. The context for testing was explained as obtaining psychological test data from the group as the basis for learning how to use the analysis of variance procedure. Laboratory work utilised teacher-made handouts (i.e., no reference was made to the textbook) and was assessed independently of the material relevant to the test. Some 178 students completed all three measures.

The study was thus a 2 (Achievement Test Form) x 3 (Hidden Figures Test Classification) factorial design, with dependent measures being achievement-test score and scores concerning beliefs about test headings. In a set of preliminary analyses of achievement scores and beliefs about headings, sex of student was not associated with any significant main effects or interactions.

RESULTS

Achievement

Achievement-test scores were analysed in a 2 (Test Form) by three (Field-Independence) analysis of variance (N = 233). The mean scores associated with this analysis are shown in

Table 1. The main effect for Test Form was not significant (F = 0.50), with achievement being very similar in the "Headings" (mean = 24.97, SD = 5.39, N = 113) and the "No Headings" (mean = 25.37, SD = 5.15, N = 120) groups. As expected, the main effect for field-independence was significant, F (2,227) = 6.52, P<0.002, with achievement scores rising across the low (mean = 23.79), medium (mean = 25.33), and high (mean = 26.73) field-independence groups. However, there was no significant interaction between test form and field-independence F (2,227) = 1.86, P = 0.16. In brief, headings did not facilitate test performance.

TABLE 1

MEAN ACHIEVEMENT SCORES AS A FUNCTION OF TEST HEADINGS AND FIELD INDEPENDENCE GROUP

		Field Independence		
Test Form		Low	Medium	High
Headings	Mean	24.11	24.18	26.89
	SD	5.87	4.66 .	4.98
No Headings	Mean	23.53	26.40	26.55
	SD	5.26	4.54	5.14

Beliefs about headings

Scores on the seven items of the questionnaire of beliefs about the effects of headings in tests were analysed in a 2 (Test Form) x 3 (Field-Independence) multivariate analysis of variance (MANOVA). The mean scores associated with this analysis are shown in Table 2. In terms of the multivariate effects, students' beliefs did not vary significantly as a function of test form, field-independence classification, or the interaction of these two variables (P >0.25). Furthermore, none of the univariate tests of main or interaction effects reached significance.

TABLE 2

MEAN BELIEF SCORES (AND STANDARD DEVIATIONS) CONCERNING TEST HEADING EFFECTS AS A FUNCTION OF TEST HEADINGS AND FIELD INDEPENDENCE GROUP

	Headings			No Headings		
	Field Independence					
	Low	Medium	High	Low	Medium	High
Will increase score	1.35 (0.49)	1.52 (0.51)	1.36 (0.49)	1.42 (0.55)	1.39 (0.50)	1.29 (0.46)
Will help understanding	1.08 (0.27)	1.33 (0.48)	1.21 (0.42)	1.56 (0.44)	1.23 (0.43)	1.11 (0.32)
Will help remembering	1.19 (0.40)	1.19 (0.40)	1.21 (0.42)	1.21 (0.53)	1.26 (0.44)	1.07 (0.26)
Will reduce anxiety	2.11 (0.52)	2.41 (0.50)	2.18 (0.39)	2.13 (0.47)	2.10 (0.47)	2.11 (0.42)
Advantage good student	1.35 (0.49)	1.48 (0.51)	1.50 (0.51)	1.37 (0.49)	1.45 (0.51)	1.43 (0.50)
Advantage poor student	1.42 (0.50)	1.59 (0.50)	1.57 (0.57)	1.66 (0.53)	1.58 (0.56)	1.43 (0.50)
Should be used	1.15 (0.38)	1.26 (0.45)	1.21 (0.42)	1.32 (0.47)	1.39 (0.62)	1.25 (0.44)

NOTE: Belief statements are shown in brief positive form

In order to determine whether the effects of headings were perceived as more due to cognitive or affective factors, the mean score on the combined "help understanding" and "help remembering" items was compared with the score on the "reduce anxiety" item for all students who claimed that headings would result in an increase in their test score (N = 125). A preliminary 2 (Test Form) x 3 (Field-Independence) MANOVA was performed on the two belief measures, resulting in no significant effects. Mean belief scores were, however, lower for the cognitive factor (mean = 1.06, SD = 0.19) than the affective factor (mean = 1.73, SD = 0.45) in affecting test achievement, t (124) = 14.7, $P<0.001$, indicating that students believed that the facilitative effect of headings is more a function of cognitive processes than of a reduction in anxiety.

To check on the relationship between beliefs and performance, achievement scores were analysed in a 2 (Test Form) x 3 (Field-Independence) analysis of variance for those students who believed that headings would increase their test score. There was no significant effect of test headings, or its interaction with field-independence, on achievement ($Fs<1.00$). The significant effect of field-independence ($P<0.05$) paralleled that for the major analysis of test achievement noted above. Thus, belief in the facilitative effects of headings was not associated with a corresponding increase in actual test performance.

To determine whether students believed that the effects of headings would vary as a function of the adequacy of students' preparation for the test, scores on item 5 (the effect on adequately prepared students) and item 6 (the effect on students not adequately prepared) were compared. As noted earlier, the univariate effects associated with the MANOVA failed to show that student beliefs were associated with test form, field-independence, or their interaction for either of these items. Students were found to believe that effects of headings would be of greater advantage to students who had studied adequately for the test (mean = 1.42, SD = 0.49) than for those who had not studied adequately (mean = 1.55, SD = 0.54), t (204) = 2.57, $P<0.05$.

To check on the relationship between these beliefs and actual test performance, a separate 2 (Test Form) x 3 (Field-Independence) analysis of variance of achievement scores was carried out for the top 96 scores (scores of 27 or better) and the bottom 83 scores (scores of 23 or less). In neither analysis was the test-heading effect or the interaction effect significant. In other words, test headings were of no benefit to either high scorers or low scorers in this study.

Although student beliefs were unaffected by the test version and field-independence in the multiple classification analysis, it is clear from Table 2 and the above analyses that students varied in their beliefs about the effects of headings in tests. A more direct interpretation of the mean differences shown in Table 2 may be apparent in the proportions of students' responses to the belief statements. The majority of students believed that headings would increase their understanding of the test items (81 per cent), and help them remember things that they had learned which were relevant to the test (81 per cent). A somewhat lower proportion of students (61 per cent) believed that the use of headings in multiple-choice tests would lead to an improvement in their test scores, while the remainder believed that headings would make no difference to their test scores. A minority of students (21 per cent) believed that headings would help reduce test anxiety, while 75 per cent of students believed that their test anxiety would be unaffected by headings. Some 58 per cent of students believed that headings would be an advantage for students who had studied adequately for the test, while 47 per cent felt that headings would be an advantage for students who had not studied adequately for the test. Finally, in relation to the question of whether students feel that headings should be used in multiple-choice tests, it should be noted that some 73 per cent of students endorsed the use of headings, while only 1 per cent opposed their use.

DISCUSSION

The results of this study indicate that although the use of headings in multiple-choice tests has little apparent effect on achievement, and appears unrelated to psychological differentiation in cognitive style, the majority of students believe that headings serve important cognitive functions and that headings should be used to designate sections of such tests. Although this is the first study to investigate the effect of headings in a multiple-choice examination, the results complement the growing research literature on the use of headings as an instructional aid in prose learning. In particular, it appears that the influence of headings in providing cues to facilitate performance (Brooks *et al.*, 1983; Wilhite, 1988, 1989) may operate at the time that the learning material is being encoded rather than at the time of retrieval during the test.

It should be noted, however, that there are fundamental contextual differences between "within-prose-headings" and "within-test-headings" tasks. In the former situation the student is concerned with facilitating the selection and organisation of a large amount of information for later retrieval. The research of Wilhite (1988, 1989) suggests that headings do facilitate such processes, at least for students with high prior knowledge tested under experimental conditions in which they are given a passage on a single topic, a fixed and supervised study time, and immediate testing. But the effect on performance of headings provided within tests is more likely to result from facilitating the search process by activating relevant (learned) prior knowledge within domains of the material studied. This effect should be enhanced if the same headings are used at both time of encoding and time of retrieval.

In tests such as that used in the current study, the use of a chapter heading from the textbook serves to activate a particular domain of knowledge in readiness to retrieve and process information from a relatively diverse range of topics within that domain. For example, the section heading "Motivation and Learning: Personality Factors" (Chapter 14, Gage and Berliner, 1988) excludes other domains relevant to the test (e.g. formulating and classifying educational objectives, conceptions and measurement of intelligence, the cognitive processing of information), but may encompass test items covering state and trait variables, locus of control, underachievement, achievement motivation, and attribution training. In the research on headings in prose, it is likely that each one of these areas would have its own heading. In so far as the learner's task is to embed new information in existing prior knowledge during learning, and then to disembed the information during time of testing, it is likely that performance will be most greatly facilitated where headings are quite specific. Given that the current study used headings directed at macro-knowledge structures, it may be that headings can only facilitate performance if they provide access to the point at which specific information is linked to the knowledge structure. This suggestion is reinforced by the failure of headings to provide an organisational structure for students low in field independence. A logical extension of this would be that each item would require its own specific heading — a procedure which would have very limited practical application.

Of relevance to the current study, investigation of the effects of headings in the retrieval process in regular educational contexts is compromised by variability both across students and within students in study behaviour (e.g., engaged study-time, amount of material covered, study techniques adopted). This results in differences in prior knowledge across the domains being assessed by the test. It may be, for example, that headings in tests facilitate performance in situations where the amount and diversity of material to be learned is more limited than in the current study, thus ensuring that the encoding phase has been adequate for all participants.

The lack of facilitation of test performance in the current study should not necessarily be regarded as an argument against the use of headings. Students were clear in their endorsement of headings and these beliefs may be encouragement enough for teachers to use them. In this respect it is interesting to note the reflective nature of students' metacognitive judgments about their memory processes. Over 80 per cent of the written comments made by students concerned some cognitive feature of the effects of headings. A number of students suggested that headings perform a knowledge-sensitising function prior to answering particular items; in this sense they help "put the questions in context", "focus attention," "orientate thinking," "direct thoughts in a particular direction," and "structure thinking". Many students made reference to memory structures or to content in memory, for example: "help you predict what is to happen — and help you relate the information that you have already studied"; "trigger some memory of learned facts"; "speed up comprehension and serve as memory joggers" and "help with retrieval, as information is learned in blocks under headings and thus it is probably easier to retrieve it that way". A few comments suggested that while headings would not affect performance, they may still serve a useful function, such as "I could see at a glance the sections I wished to answer first," or to "allow the testee to do the sections in order of confidence," or to "reduce the ambiguity of some test questions by making it clear what topic they are assessing". Although almost all comments argued for a facilitative effect of headings, a few argued for a negative effect, for example, "they may tend to interrupt your thought pattern when you are in the habit of reading questions," and "if you have not studied adequately, it may be disconcerting to have a heading for a chapter you have not read".

In conclusion, the use of test headings appears justified on the grounds of student popularity, though its potential to increase achievement remains to be determined in studies varying the specificity of the headings to the knowledge being assessed by the test items, and the prior knowledge students have when they enter the test situation. The importance of this prior knowledge is evident in one student's comment which is appealing for its clarity and simplicity: "It's not really the headings that make a difference, it's the questions."

Correspondence and requests for reprints should be addressed to Michael A. R. Townsend, Education Department, University of Auckland, New Zealand.

REFERENCES

Anderson, R. C., Spiro, R. J., and Anderson, M. C. (1978). Schemata as scaffolding for the representation of information in connected discourse. *Am. educ. Res. J.*, 15, 433-440.

Brooks, L. W., Dansereau, D. F., Spurlin, J. E., and Holley, C. D. (1983). Effects of headings on text processing. *J. educ. Psychol.*, 75, 292-302.

Dee-Lucas, D., and Di Vesta, F. J. (1980). Learner-generated organizational aids: effects on learning from text. *J. educ. Psychol.*, 72, 304-311.

Dodd, D. K., and Leal, L. (1988). Answer justification: removing the "trick" from multiple-choice questions. *Psychology of Teaching*, 15, 37-38.

Flaugher, R. L., Melton, R. S., and Meyers, C. T. (1968). Item rearrangement under typical test conditions. *Educ. psychol Measur.*, 28, 813-824.

Frank, B. M. (1984). Effect of field independence-dependence and study technique on learning from a lecture. *Am educ. Res. J.*, 21, 669-678.

French, J. W., Eckstrom, R. B., and Price, L. A. (1969). *Kit of Reference Tests for Cognitive Factors*. Princeton, N.J.: Educational Testing Service.

Gage, N. L., and Berliner, D. C. (1988). *Educational Psychology* (4th ed.). Boston: Houghton Mifflin.

Hartley, J., and Trueman, M. (1983). The effects of headings in text on recall, search and retrieval. *Br. J. educ. Psychol.*, 53, 205-214.

Hartley, J., and Trueman, M. (1985). A research strategy for text designers: the role of headings. *Instruct. Sc.*, 14, 99-156.

Kardash, C. M., Lukowski, L., and Bentmann, L. (1988). Effects of cognitive style and immediate testing on learning from a lecture. *J. educ. Res.*, 81, 360-364.

Kiewra, K. A., and Frank, B. M. (1988). The encoding and external-storage effects of personal lecture notes, skeletal notes, and detailed notes for field-independent and field-dependent learners. *J. educ. Res.*, 81, 143-148.

Skinner, N. F. (1983). Switching answers on multiple-choice questions: shrewdness or shibboleth? *Teaching of Psychology*, 10. 220-222.

Spiro, R. J., and Tirre, W. C. (1980). Individual differences in schema utilization during discourse processing. *J. educ. Psychol.*, 72, 204-208.

Br. J. educ. Psychol., 60, 161-170, 1990

RESPONSE FEEDBACK, CERTITUDE AND LEARNING FROM TEXT

By Raymond W. Kulhavy, William A. Stock, Nancy E. Thornton, Kristen S. Winston, and John T. Behrens
(Arizona State University, U.S.A.)

Summary. Based upon the premise that persons modify their study behaviours in order to maintain their certainty of responding correctly, a study was conducted with 40 undergraduate volunteers examining the effects of different presentation rates and feedback context on response confidence and correct responding. Significant effects in the anticipated direction were found for presentation rate but not for the context of feedback.

INTRODUCTION

Information processing frameworks have freed instructional feedback from its early association with operant psychology, and, as a consequence, such feedback is infrequently assumed to have reinforcing properties. Research on the effects of instructional feedback shows that one of its important functions is the correction of errors (e.g. Kulhavy, 1977). However, there are also data which suggest that simple error correction is an inadequate explanation for important events which occur as a result of being exposed to a post-response knowledge standard (Kulhavy *et al.*, 1976, 1979). It is our view that a key to understanding how feedback influences instructional behaviour lies in a learner's perceptions of tasks, feedback, and tests, rather than in the objective content of feedback. In a previous paper we have defined a learner's perceptions of the correctness of their response as their degree of certitude in that response (Kulhavy and Stock, 1989). In this case we accepted the O.E.D. (1933) definition of certitude as, "Subjective certainty; the state of being certain or sure of anything". Although certitude judgments about instructional responses are not perfect, they do provide useful information about the cognitive status of the response in question.

When certitude is operationalised as a rating of correct response confidence, Kulhavy *et al.* (1976, 1979) found a consistent relation between response ratings and learning when questions dealt with a previously studied text segment. Results from these studies indicate post-response study time is directly related to response confidence for persons making an incorrect response and inversely related to confidence for those making a correct response. Individuals choosing incorrect responses with high confidence are more likely to correct such errors than those choosing incorrect responses with low confidence. Such results suggest that response certitude has an important role in the acquisition of knowledge from text. This article presents a rationale for the inclusion of certitude in models of text processing, and reports a study designed to evaluate this rationale.

We begin with the basic assumption that an instructional task involves the achievement of some performance criterion, simply because this production is a way of showing that a planned increment of learning has occurred. Hence, a critical variable is the learner's perception that the criterion demand can be met after studying. Our contention is that it is the learner's perception that he will be able to perform correctly which shapes subsequent behaviour in the instructional milieu.

The student's estimate of certitude is determined by the degree of match between perceived comprehension of the criterion and the episodic/semantic prior knowledge available to the learner. We propose that learners compare their perception of the criterion demand against a set of cognitive referents consisting of both content ("Do I know enough about this material to understand it?") and experience ("Do I generally perform well on this type of task?") information. The degree of match between the perceived comprehension and the cognitive referents creates an expectation that the criterion response produced will be the correct one (e.g., the one desired by the creator of the instruction). We further argue that it is the magnitude of this expectation which determines the behaviours of the learner when feedback is presented after the instructional response.

Our model of how response expectancy influences future instructional behaviours employs concepts derived from control theory (e.g., Powers, 1978). Using the metaphor of control theory, "perceived comprehension" is a controlled quantity, and the cognitive referents to which it is compared form the basis for a reference standard. In order to simplify our presentation, we characterise perceived comprehension as a representation of the perceptions of the instructional materials and the response demands (e.g., the questions testing these materials). This representation is matched to materials known by the learner, including information in the instructional message which have become incorporated in the knowledge structure of the individual (Kulhavy and Stock, 1989).

In control theory, the operation of a comparator (here the matching of perceived comprehension with the knowledge base standard) is a continuous process, with control of processing managed by negative feedback loops. In a negative feedback loop, behaviour is directed at maintaining the value of a control variable at a specified reference value. A similar position may be held here, and one can view such a comparison as controlling behaviours involved in the task of acquiring the instructional material.

Theoretically, we view the feedback message which follows a response as being composed of the two separable components, verification and elaboration (Kulhavy and Stock, 1989). The verification component refers to the simple dichotomous judgment that the initial instructional response was right or wrong. Hence, when a student is simply told "yes" or "no", the only function served by feedback is verification. The elaboration component consists of all substantive information contained in the feedback statement, such as representing the entire item, or providing an explanation regarding why the response was right or wrong. For ease of understanding we label the verification component as (f_v) and any added elaborations as (f_l). We can now construct a notational shorthand for the two effects that feedback has on a learner:

(1) Feedback = $f_v + f_l$

In the present context, we prefer to focus on two points in the instructional sequence, the first shortly after a learner's response, and the second after the presentation of instructional feedback. Therefore, we suggest that following instruction and a response the learner is sensitive to the likelihood of a response being correct (a consequence of the knowledge monitoring comparison). The magnitude of this likelihood will range from low to high, with higher values reflecting a higher likelihood that the response is correct, according to the standards and perceptions of the learner.

With presentation of instructional feedback, the learner is informed with respect to an external standard for the response. As a consequence of a comparison between the two perceptions derived from feedback and the prior response (as well as the expectation of responding correctly), the original estimate of performance certitude is transformed into a measure of discrepancy. According to Kulhavy and Stock (1989), discrepancy (d) is a measure of error between input and reference standard which takes into account the degree of match between the initial response (r_i) and the feedback message to produce a verification estimate (f_v). The verification estimate is then combined with the initial response certitude (c) to yield a discrepancy value for the response in question. Hence, using our notational shorthand, discrepancy is defined as:

(2) $f_v \times c = d$

The scale used to distribute (d) for initially correct and error responses is developed more completely in the Appendix which follows this paper.

It is the magnitude of the discrepancy value that we believe controls subsequent studying behaviours. Here, we suggest that a simple behavioural mechanism is engaged, where the learner undertakes studying behaviours in order to learn responses which eliminate mismatches between the initial response and feedback. These subsequent study behaviours will emerge immediately. For example, we expect learners to study feedback longer when it is associated with incorrect responses. Also, if learners are given a choice to restudy instruction, they will choose to restudy more often those text frames associated with incorrect responses, and to spend more time studying those particular frames. There will also be a greater likelihood that high confidence errors will be replaced by correct responses compared to low confidence errors.

In designing the present study, we reasoned that limiting the time available to study initial instruction would have two effects. First, shorter study times would directly affect the level of certitude learners expressed in being able to answer questions over the material studied, with shorter presentation intervals associated with lower expressions of response certitude. Second, we expected that subsequent feedback study time would be affected similarly, with shorter text presentation rates being associated with longer feedback study times. Thus, in this study a between-subjects factor included three levels of text presentation, with fast, average, and slow rates determined by times available from a norming study.

Next, we were interested in whether structural changes in the feedback messages would affect feedback study time. While Kulhavy *et al.* (1985) found that feedback study times increased as a function of increasing complexity of feedback (e.g., increases in the magnitude of f_i) our intent was to examine the effects of changes which did not alter the length of the feedback messages. Therefore, we altered the context of the feedback message, as a within-subjects factor, by using either the original test item as the basis of feedback (in context), or the original test item with all original distractors replaced by new plausible alternatives which were structurally equivalent to the old. In this way we hoped to tease out the influence of minor changes in the context of (f_i) on subsequent responding.

METHOD

Design and participants

Three levels of frame reading time varied between-subjects and crossed with two levels of feedback context within-subjects. Three reading times had values of slow, average, and fast, while feedback was formed by an in vs. out of context comparison. Participants were 54 undergraduate volunteers who were randomly assigned to between-subjects cells in order of their appearance for the experiment. Data from 14 individuals were omitted when the microcomputer control program failed to transfer to diskette a single time measure on each of these persons, leaving a final sample size of 40.

Materials

Text materials consisted of 12 paragraph length frames describing the United States Navy. Frames averaged 147 words in length, and were written so that they could be read in any order without breaking the general continuity of the passage. A five-alternative, multiple-choice question was based on each frame. Both frames and questions were modified versions of materials developed and normed earlier (Kulhavy *et al.* 1985).

The text was easily comprehensible by an undergraduate population, and all test items could be answered correctly in a straightforward fashion from the information contained in the paragraph. An example paragraph and its corresponding test item are displayed below.

> The fleet that is able to carry on its own commerce and stop that of an opponent is said to be in "command of the sea" in the region where it enjoys that marked advantage. The term "command" in this case does not mean complete control, but only that sea-borne communications over certain areas are restricted. A fleet can secure command of the sea either by destroying the enemy or by denying him access to the sea area. The first of these alternatives is best because it settles the matter completely and permanently. A fleet that refuses battle and stays in port does not cease to be a threat. The act of avoiding battle while still maintaining the maximum threat value is known as the strategy of the "fleet in being" and forces the fleet in command of the sea area to be constantly on guard.
>
> Concerning command of the sea, which of the following statements is true?
>
> a. When a fleet commands a sea area, it is not actually interested in movement of the enemy ships themselves.
>
> b. The phrase "fleet in being" refers to a mobile fleet that is eager to come to battle as soon as possible.
>
> c. Only the ships of a major fleet are able to exercise command of the sea in a given area.
>
> d. Command of the sea area simply means that a fleet controls communications over that area.
>
> e. The best possible way to command a sea area is to deny an enemy access to that area.

In-context feedback items were constructed by capitalising and marking with an asterisk the correct alternative in the corresponding question, otherwise leaving the original question intact. Out-of-context feedback items were created by substituting new distractors in the respective in-context feedback item, and marking the correct alternative in the same way as in-context items. The four new distractors were constructed from information contained in the target paragraph, were of the same length as the originals, but differed entirely in their content. Each learner saw six in-context and six out-of-context feedback items, with the specific selection of feedback items separately randomised for each participant.

During the course of the study learners responded three separate times to a five-point scale which asked them to rate their confidence in their current task. Following the reading of a paragraph they were asked to rate their confidence in being able to answer a multiple-choice question over the material just read. Response options ranged from *not at all sure* (1), through *somewhat sure* (3), to *absolutely sure* (5), with points (2) and (4) not labelled. Following their initial response to a test item, and again after their response on the post-test, they were asked to rate how confident they were that the answer just given to the question was the correct one. The scale used in the later two cases had point labels identical to those described above.

A post-test consisted of the 12 items originally responded to during the text presentation. Items were printed on 8¼ x 11 inch (21.25 x 27.50 cm.) white paper and stapled into test booklets for each participant, with the order of items separately randomised for each learner.

Frame presentation rates were derived from a norming study using 20 undergraduates who read the paragraphs at their own rate. Our average presentation time was the overall mean per frame time (M = 52 sec.). Our slow and fast presentation times were one pooled standard deviation (SD = 16 sec.), above (68 sec.) and below (36 sec.) the average time, respectively.

Procedure

An IBM BASIC program controlled presentation of materials in the study, with individual randomisations of text frames and context items for each learner. As learners entered the room they were randomly assigned to one of the between-subjects conditions and seated in front of a microcomputer. A monitor told them that they would be reading a series of paragraphs about the Navy, answering a question over each paragraph, and rating their confidence in their ability to answer questions from the text, and in how correct they thought their response was to each test item. A monitor also provided an introduction to use of the terminal, answered procedural questions, and started the program.

All participants completed a practice sequence identical in form to the presentation of experimental materials. During the practice session learners read a paragraph on the screen, rated confidence, responded to a test item, rated confidence, and practised the reread option described below. Subsequently, they worked through the stimulus portion of the experiment. For each of the 12 paragraphs, learners studied the text for their allotted time, rated their confidence in their ability to answer questions over the material, responded to test item, rated their response confidence, studied the feedback (in or out of context), and were finally given the option of rereading the paragraph prior to moving to a new frame. Each learner saw six in-context and six out-of-context feedback items.

Each element of the steps above was presented as a separate screen, with the material for a screen presented simultaneously, rather than line by line. Response choices, confidence ratings, and reread selections were recorded on-line by the computer. Latencies were also recorded for text reading, item responding, feedback study, and reread selections. All responses were automatically written to diskette. When the stimulus portion of the experiment was completed, learners were given a copy of the post-test booklet and given unlimited time to answer the original 12 items. The last page of the booklet contained a general questionnaire asking participants about what strategies they used in completing the experimental tasks.

RESULTS

Responses to multiple choice questions were scored for correctness; number of reread options were summed for correct and incorrect responses items within context classifications; and question, feedback and text restudy times were converted to log times according to procedures recommended by Winer (1971, p. 401). Generally, two planned comparisons were used throughout to decompose effects associated with presentation rate; one contrast compared the fast presentation group (N = 14) to the average presentation group (N = 12), while the second contrast compared the combined fast and average presentation groups against the slow presentation group (N = 14). Examined first was correct responding as a function of text presentation rate, context of feedback, and occasion (items immediately following instruction vs. on the post-test).

Correct responses

Table 1 provides summary statistics for correct responses.

TABLE 1

MEAN CORRECT INSTRUCTIONAL AND POST-TEST RESPONSES FOR FAST, AVERAGE, AND SLOW PRESENTATION GROUPS

	Presentation Group					
	Fast		Average		Slow	
Measure	Mean	SD	Mean	SD	Mean	SD
Instructional Corrects	5.93	2.64	8.17	2.95	9.36	2.41
Post-test Corrects	11.07	1.38	9.83	3.04	11.21	1.53
N =		14		12		14

Analysis of variance confirmed that overall recall for the fast and average rates of presentation group did not differ ($F(_{1,37}) = 1.62$), while these groups combined differed significantly from the slow presentation group ($F(_{1,37}) = 4.87$, $P<0.05$). Levels of correct responses to instructional items were significantly lower than to post-test items ($F(_{1,37}) = 66.02$, $P<0.01$). There were also significant interactions involving each of the one degree of freedom comparisons and time of test; with one test indicating that the group receiving the fastest presentation rate of text scored lower on the instructional items and higher on the post-test items than the group receiving the average rate of presentation ($F(_{1,37}) = 15.52$, $P<0.01$); and the second test indicating that the group receiving the slowest rate of presentation scored higher on both instructional and post-test items than the combined group, although the difference was about three times larger on instructional items ($F(_{1,37}) = 4.32$, $P<0.05$). Correct responding did not differ as a function of the context of feedback or any other effect remaining in the model. Because context was not a significant factor in the analysis of correct responses, further analyses were collapsed over this dimension.

Study times and rereading

Table 2 contains summary statistics for question and feedback study times. In this table, times have been reported in both log seconds and seconds to facilitate the reader's examination of the table.

TABLE 2

MEAN STUDY TIME FOR QUESTIONS AND FEEDBACK FOR FAST, AVERAGE, AND SLOW PRESENTATION GROUPS

Type		Presentation Group					
		Fast		Average		Slow	
Measure		Mean	SD	Mean	SD	Mean	SD
Instructional items							
Corrects							
	log seconds	1.48	0.12	1.53	0.11	1.48	0.11
	seconds	32.21	8.75	36.85	9.39	32.72	9.23
Incorrects							
	log seconds	1.55	0.12	1.60	0.13	1.56	0.15
	seconds	38.81	11.07	43.55	10.68	40.85	15.12
Feedback items							
Corrects							
	log seconds	0.73	0.18	0.80	0.24	0.55	0.28
	seconds	6.55	2.36	8.41	5.09	5.06	3.87
Incorrects							
	log seconds	0.91	0.13	1.04	0.27	0.83	0.21
	seconds	9.49	3.67	14.38	8.17	8.33	3.53

NOTE: Sample sizes for the fast, average, and slow presentation group means were 14, 12, and 13, respectively. One subject in the slow presentation group was dropped from the analyses corresponding to this table because of an absence of incorrect responses on which to compute study times.

Turning to an examination of time spent studying and responding to questions, analysis of variance did not reveal a significant effect for rate of presentation, but did for correctness of response ($F(_{1,36}) = 13.64$, $P<0.01$), with subjects spending less time responding to items they answer correctly than to those they answer incorrectly (about 34 seconds versus 41 seconds).

Feedback study times were found to vary as a function of presentation rate, with the combined fast and average group studying feedback significantly longer than the slow group, ($F(_{1,36}) = 7.14$, $P<0.02$), but not differing from one another. Correct responses were followed by significantly shorter feedback study times than were incorrect responses, ($F(_{1,36}) = 45.92$, $P<0.01$).

Analysed next were text reread choices of subjects. Table 3 provides a summary of the proportion of items on which learners chose to reread text frames.

TABLE 3

Proportion of Items on which Subjects chose to Reread the Text for Fast, Average, and Slow Presentation Groups

	Presentation Group		
Measure	Fast	Average	Slow
Instructional Corrects	0.02	0.13	0.02
Post-test Corrects	0.32	0.23	0.16
N =	14	12	14

Analysis of variance revealed a significant effect for the comparison of the combined fast and average groups relative to the slow presentation group ($F(_{1,37}) = 3.91, P = 0.05$), while the fast and average groups did not differ. With a combined proportion of 0.18, individuals in the fast and average groups chose to reread about one in six text frames, a rate about three times more frequent than individuals in the slow presentation group, who reread about one in 20 text frames. The number of reread choices also varied as a function of the correctness of response ($F(_{1,37}) = 4.40, P<0.05$), with people choosing to reread text frames on which they made errors on the immediate test items about five times more frequently than text frames on which they responded correctly (proportions of 0.29 and 0.06, respectively). No other effects were significant.

No differences in text rereading time were found among groups receiving different presentation rates, but there was a significant difference for correctness of response ($F(_{1,37}) = 28.25, P<0.01$). Learners spent less time rereading text associated with correct responses than with incorrect responses, 1.63 and 6.31 seconds, respectively.

Response confidence

Expressions of confidence made immediately before responding (text), immediately after responding (item), and after post-test responding (post-test) were averaged across the 12 text frames and analysed in separate one-way analyses of variance using planned comparisons. The findings of each analysis were identical and mirrored those for correct responding, with groups receiving fast and average rates of presentation not differing significantly, while their combined mean was significantly lower than that for the group receiving the slow rate of presentation ($F(_{1,37}) = 11.34$. $P<0.01$ for text confidence; $F(_{1,37}) = 6.87, P<0.02$ for item confidence; and $F(_{1,37}) = 4.84, P<0.05$ for post-test confidence).

For the fast, average and slow groups, respectively, mean response confidences were 3.14, 3.12, and 3.91 for text confidence, 3.14, 3.32, and 3.81 for item confidence, and 4.45, 4.29, and 4.74 for post-test confidence. We note two phenomena here. First, the crossover in mean confidence ratings for the fast and average group reflects a corresponding crossover in immediate and delayed correct performance. Second, there is some convergence, but no elevation, of confidence ratings after exposure to the questions, followed by further convergence and elevation of confidence after exposure to feedback.

Examined last were the effects of response confidence, the experimental factors, and correctness of response on feedback study time. We used two complementary analyses.

Preliminary to these analyses we examined the distributions of correct and incorrect responses as a function of response confidence. Errors were positively skewed (few high confidence errors), while corrects were negatively skewed (few low confidence corrects). Thus, we recoded response confidence into low (1 or 2 on the confidence scale), average (3), and high (4 or 5), and, using analysis of variance, examined log feedback study time as a function of response confidence, presentation rate, context, and correctness of response.

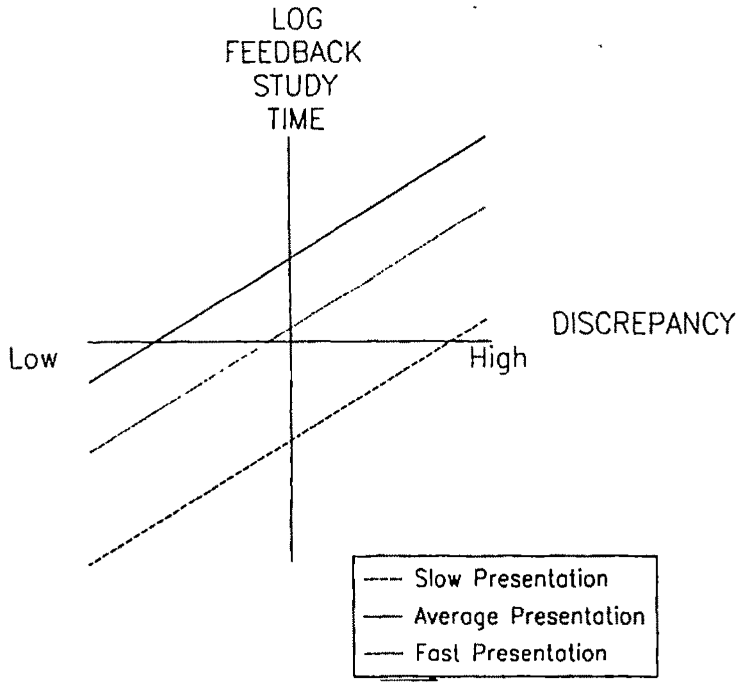
Log feedback decreased systematically as response confidence increased ($F(_{2,468}) = 3.45$, $P<0.05$, and was significantly longer following error responses ($F(_{1,468}) = 31.77$, $P<0.01$). These findings occurred in the context of a significant response confidence by correctness of response interaction ($F(_{2,468}) = 5.86$, $P<0.01$), where log feedback time systematically increased as confidence increased for incorrect responses, and systematically decreased as confidence increased for correct responses. No other effects were significant.

In the complementary analysis we transformed confidences by first recoding as above and by calculating discrepancy values according to equation (2) above (see Appendix). Additionally, context of feedback was dropped from consideration. Hence, we were able to examine, in an analysis of covariance, log feedback time as a function of discrepancy (covariate) and presentation rate (experimental factor).

A test of parallelism was not significant, indicating a pooled estimate of the slope was adequate across groups. The estimated slope was significantly different from zero ($F(_{1,474}) = 67.69$, $P<0.01$), with a value of 0.35, indicating a consistent increase in log feedback study time as a function of increasing discrepancy. Adjusted group means were significantly different ($F(_{2,474}) = 13.32$, $P<0.01$).

FIGURE 1

STANDARDISED REGRESSION LINES RELATING LOG FEEDBACK STUDY TIME TO DISCREPANCY VALUES FOR THE FAST, AVERAGE, SLOW PRESENTATION GROUPS.



NOTE: All values are z-units

In Figure 1, a plot of the group regression lines on z-score axes, it can be seen that the intercepts of the average (0.25), fast (0.05), and slow presentation (-0.27) are fairly evenly separated. In examining Figure 1, remember that these intercepts only indicate relative position because of the use of z-plots, and that times are still log functions of actual study time. Despite these technical considerations, it is clear that time devoted to studying feedback is an increasing function of discrepancy, consistent in nature across experimental conditions. Hence, although intercepts changed as a function of stimulus conditions (reading time), the slopes remain essentially the same across conditions.

DISCUSSION

The manipulation of feedback content had little effect on our results, and there seems to be no reason to discuss context further, except to make two points. First, the previous finding that feedback study time increased as a function of the amount of information in the feedback statement appears to be due to the not surprising fact that more information takes longer to read (Kulhavy *et al.*, 1985). In the present case, when amount of information was held constant, changes in the content of the item alternatives had no important effect. Second, the lack of differences due to the context variable suggests that students concern themselves primarily with using feedback to verify their instructional response, rather than inspecting wrong alternatives for additional information.

Correct responding

That presentation rates were related to correct responding in the anticipated direction is not surprising, for a generally accepted positive relation exists between time on task and correct performance. However, a network of relations linking study to response confidence, subsequently to correct performance, while also explicating the relation of response confidence to further study behaviours for both correct and incorrect responses is less generally accepted. In this context, it is important to note that levels of correct responding reflected exactly levels of immediate expressions of response confidence. That fast and average rates were not different overall is mitigated by the presence of a significant interaction involving these two groups and the test occasion. Also attributable to the presentation and study of feedback are the increases in correct responding from the instructional to the post-test, as well as the convergence of scores for the three groups. While scores on the post-test are approaching the ceiling of the scale, the slow representation group did gain about one scale unit from immediate to delay occasions, and the increases in scores followed exactly differences in time spent studying feedback, to which we now turn our attention.

Study times

Feedback study times varied in manner consistent with correct performance: groups who performed less well studied feedback longer. Again, the principal difference involved the comparison of the combined faster presentation groups with the slow presentation group. The important point here is that differences in feedback study time corresponded directly to differences in expressions of post-response confidence, with the latter measured prior to a subject verifying their response. Certitude judgments seem to be a key in this network.

Response confidence

Immediate expressions of response confidence are a direct consequence of presentation rate, and precede corresponding differences in correct responding. Subsequent to responding, but prior to verification, post-response confidence is related not only to presentation rate but also to whether or not the response turns out to be correct. Differences in post-response confidence then precede corresponding differences in feedback study time, which are subsequently followed by corresponding increases in correct performance on the post-test occasion. Here, performance tends to converge, as do final expressions of response confidence.

Further, an item level examination of the relation between response confidence and feedback study indicates that subjects differentially allocate their study time. We interpret this allocation to be a function of the calculated value of discrepancy, and found the allocation to have a consistent functional relation across conditions.

Conclusions

The data reported in this paper are in line with the feedback certitude model proposed by Kulhavy and Stock (1989), and support the resulting predictions which we articulated in the introductory statement. The variable of response certitude, operationalised as a confidence judgment, has been the centre of our attention in this particular study. By and large, the results meet our expectations for the kinds of relationships that should exist between certitude and both feedback study time and correction of initial errors.

We were especially pleased with the form of the three functions shown in Figure 1. We initially argued that the amount of time a student spends working to correct errors at feedback is a strong positive function of discrepancy in the processing system. These results are exactly the ones depicted in the plot, and they may indicate that it is the intercept values, but not the slopes, which are sensitive to manipulations of instructional task demands. Given this state of affairs, we might then predict that the slope values would change for different groups depending on manipulations that affect variables during feedback. Such reasoning leads to some interesting theoretical possibilities, not the least of which is that the intercept/slope values in Figure 1 represent relatively stable parameters of the learner's instructional processing system. If this conclusion turns out to be a valid one, then research within the framework of the model should, indeed, be profitable in the future.

ACKNOWLEDGEMENT.—The research reported here was supported, in part, by an Arts, Social Sciences, and Humanities Grant from the Office of the Vice-President for Research, Arizona State University.

Correspondence and requests for reprints should be addressed to Professor Raymond Kulhavy, College of Education, Arizona State University, Tempe, Arizona 85287-0611, U.S.A.

REFERENCES

KULHAVY, R. W. (1977). Feedback in written instruction. *Rev. educ. Res.*, 47, 211-232.

KULHAVY, R. W., AND STOCK, W. A. (1989). Feedback in written instruction: the place of response certitude. *Rev. educ. Psychol.*, 1, 279-308.

KULHAVY, R. W., YEKOVICH, F. R., AND DYER, J. W. (1976). Feedback and response confidence. *J. educ. Psychol.*, 68, 522-528.

KULHAVY, R. W., YEKOVICH, F. R., DYER, J. W. (1979). Feedback and content review in programmed instruction. *Contemp. educ. Psychol.*, 4, 91-98.

KULHAVY, R. W., WHITE, M. T., TOPP, B. W., CHAN, A. L., AND ADAMS, J. (1985). Feedback complexity and corrective efficiency. *Contemp. educ. Psychol.*, 10, 285-291.

POWERS, W. T. (1978). Quantitative analysis of purposive systems: some spadework at the foundations of scientific psychology. *Psychol. Rev.*, 85, 417-435.

WINER, B. J. (1971). *Statistical Principles in Experimental Design*. New York: McGraw-Hill.

APPENDIX: DISCREPANCY CALCULATIONS

The discrepancy measure (d) portrayed on the abscissa in Figure 1 was developed by Kulhavy and Stock (1989), and is calculated using equation (2) above, where:

$$d = f_v \times c$$

In this formulation, the variable f_v is set equal to $(-1)^m$, where ($m = 0$) for initial error responses, and ($m = 1$) for initial corrects. This term has the effect of assigning an algebraic sign to (d), where $[(-1)^0 = +1]$ for errors, and $[(-1)^1 = -1]$ for correct responses.

The variable (c) represents a certitude estimate produced by a learner. In the studies reported, subjects generated certitudes by rating their confidence on a five-point Likert-type scale. Under these conditions, the scale of (d) runs from (-5) to (+5), with initial correct responses represented in Quadrant IV, and errors in Quadrant I.

Br. J. educ. Psychol., 60, 171-179, 1990

FUTURE LIVES OF TRUANTS: FAMILY FORMATION AND HEALTH-RELATED BEHAVIOUR

By Angelika Hibbett*
(*Employment Department, London*)
and Ken Fogelman
(*School of Education, University of Leicester*)

Summary. This paper examines early adult outcomes of truancy regarding family formation and health-related behaviour, and is based on data collected in the National Child Development Study, which surveyed all people in Great Britain born in the week 3-9 March 1958, at birth, and at ages 7, 11, 16 and 23. Truants compared to non-truants were found to be relatively more likely to have experienced marital breakdown by the age of 23. They had more children, and their age at the birth of their first child was younger. Truants differed little from non-truants regarding their drinking habits, but were more prone to being heavy smokers. They also showed an increased likelihood of depression. Differences remained after controlling for social background, prior educational attainment, school attendance and qualifications obtained. The authors conclude that truancy appears to be associated with subsequent marital and psychological problems in early adulthood.

INTRODUCTION

This paper is the second to describe the findings of an investigation based on the National Child Development Study (NCDS) designed to examine the relationships of truancy from school with a wide range of outcomes in early adulthood. A previously published paper (Hibbett *et al.*, 1990) examined occupational outcomes of truancy. In the present paper we shall discuss social and personal outcomes associated with truancy: family formation patterns and some aspects of health-related behaviour, namely smoking and drinking habits, and an indicator of depression.

Much of the previous research into truancy has investigated its causes and related behavioural problems in the children concerned. The present project aimed at examining a different aspect of this subject: the outcomes in later life associated with truancy. The point of looking at outcomes is to understand causes. Competing explanations for truancy have focused on personal characteristics and maladjustment, on work-orientation and on the characteristics of schools and their curricula. Each of these explanations generates a range of hypotheses concerning the expected relationship between truancy and experiences after school. If it is correct that truancy is an indication of a more general maladjustment (Tyerman, 1968), then it can be hypothesised that there would be continued evidence of this in, for example, family formation or on indices of mental health in early adulthood. If, on the other hand, truancy indicates a move towards the values of work and away from the values of the school (Haigh, 1976), then those who had truanted from school would not be expected to have more difficulties than their peers in later life.

There have been few studies of family formation and health-related behaviour of truants. In a study of causes and effects of persistent job changing, Cherry (1976) established that low attendance was related to an increased likelihood of broken marriages, illegitimate births and psychiatric problems. In a subsequent longitudinal survey from age 8 to age 21 of 411 male respondents in a working class area in London, Farrington (1980) found that truants tended to have a markedly antisocial or deviant life style after leaving school, and that they were more likely to be involved in antisocial activities. Truancy was also found to be significantly related to juvenile delinquency and violence. Truants were more sexually active, and they had an increased likelihood of being heavy smokers, but there was comparatively little difference in their likelihood of heavy drinking.

---

* This research was carried out when Dr Hibbett was employed at the Social Statistics Research Unit, City University, from 1986 to 1988.

These findings were confirmed by Robins and Ratcliffe (1980), who investigated 235 black schoolboys from St Louis, born between 1930 and 1934. High school absenteeism was strongly related to alcohol abuse, criminality, marital problems and parental irresponsibility (i.e., fathering an illegitimate child, or failing to support legitimate children), and, to a lesser degree, to anxiety and depressive disorders and psychiatric problems. These outcomes were predicted by absenteeism even when controlling for the effects of non-graduation and adolescent deviance. An interesting social outcome was a tendency towards "assortative mating", i.e., marrying absentee girls, and producing absentee children. More recently, in a study of the relationship between children's and parental smoking habits, and absence from school, Charlton and Blair (1989) found that there is a positive correlation between smoking and absence from school. Children who smoked were more likely to be absent from school, and parental smoking increased that likelihood still further.

In a study of psychological outcomes, Tyrer and Tyrer (1974) found school refusal to be significantly correlated to adult neurotic illness, i.e., depressive and anxiety disorders, particularly for female respondents. Similar findings were reported more recently by Watters (1989), and a study based on the Medical Research Council's National Survey of Health and Development (Rodgers, 1990), which established a tendency towards psychological disturbances associated with truancy, and also confirmed the previously observed differential between male and female respondents.

While there is thus some evidence to suggest that truancy is associated with juvenile delinquency, increased smoking, and adult neurotic illness like depression, research on family formation appears to have been neglected so far. Robins and Ratcliffe (1980) investigated some of these aspects, but it has to be seen whether these American findings are applicable here. Certainly, regarding other aspects, namely subsequent earnings of truants, their findings were contradicted by British research (Cherry, 1976; Hibbett *et al.*, 1990). The present paper therefore examines a variety of social outcomes of truancy, utilising the data contained in the National Child Development Study (NCDS).

METHOD

The National Child Development Study, which surveyed some 15,000 respondents born between 3-9 March, 1958, at ages 7, 11, 16 and 23, collected an extensive amount of data on children's health, education, social background and family life. Later, exam results were included, as well as details of their occupations and careers, marriage and parenthood, and physical and mental health and illness. Response rates have been about 90 per cent during the school years and 76 per cent at age 23. The breadth of data available and the longitudinal structure of the NCDS make it possible to examine the relationship between reported truancy from school, based on information obtained during the school years, and a number of outcomes explored by interview at age 23. For further information on the NCDS see Shepherd (1985).

Based on teachers' ratings of individual pupils' truancy, four groups were identified: those said to be truanting at 11 but not at 16; those said to be truanting at 11 and 16; those said to be truanting at 16 but not at 11; and those not said to be truanting at either age. Table 1 presents the numbers of cases in each category. The first group was too small to be analysed, and the second could only be included in the initial bivariate analysis, which showed that there was a regular correlation between truancy and virtually all the outcomes examined, with truants at 16 only in an intermediate position between truants at 11 and 16 and non-truants. Tables for these findings and a fuller discussion are presented in Hibbett (1987).

After exploring the simple associations of truancy and the outcome variables, using bivariate statistics, multivariate analyses were carried out to establish which association of truancy with differences in outcomes remained after accounting for the influence of other factors, related to truancy and which could affect outcomes in early adulthood. For each outcome variable a sequence of analyses was carried out, each stage designed to account for the influence of other factors on differences between truants and non-truants.

TABLE 1

NUMBER OF RESPONDENTS

Number of Cases Column %	Boys	Girls	Total
Truanting at 11 only	26	15	41
	0.5	0.3	0.4
Truanting at 11 and 16	51	11	62
	0.9	0.2	0.6
Truanting at 16 only	1179	888	2067
	21.7	17.1	19.4
Not truanting at 11 and 16	4178	4292	8470
	76.9	82.4	79.6
Column Total	5434	5206	10640

For each outcome, the initial contrast presents the simple differences between truants at 16 and non-truants. It had been shown in earlier work based on the NCDS (Fogelman and Richardson, 1974) that truants were relatively more likely to be from families with fathers in manual occupations, that they were not equally distributed across the country, and that they showed poorer educational ability and attainment compared to other respondents.

In the second stage of the analyses, the influence of social background factors (social class of father's occupation, region of residence, number of children in the household) is accounted for, i.e., the remaining contrast between the two groups presents differences between truants and non-truants of similar social background.

Likewise in the third stage, educational attainment prior to secondary schooling (General Ability Test score, Reading Comprehension Test score, Mathematics Test score, each at age 11) is accounted for. Truants are also likely to have lower-than-average school attendance, and lower attendance for whatever reason is in itself related to lower attainment (Fogelman, 1978). Therefore the fourth stage of the multivariate analysis concerns the respondents' attendance rate at 15, to establish the degree to which differences in outcomes between truants and non-truants are due to poor attendance, which may be due to, for example, chronic illness rather than truancy.

Finally, a previous study focusing on occupational outcomes (Gray *et al.*, 1980) found that truants' difficulties in this respect were wholly explained by their poorer examination performance, i.e., that truants and non-truants with similarly poor qualifications experienced similar problems. The fifth and final stage of the multivariate analyses therefore controls for qualifications obtained between ages 16 and 23, to determine to what extent different outcomes are due to lack of qualifications. Although the role of qualifications appears less likely to be relevant to the outcomes presented here than to those concerned with labour market experiences, these findings are presented for completeness and comparability with our earlier paper.

With this design it was possible to identify the contribution made by each factor or set of factors to the relationship between truancy and the outcome variable. By accounting for the effects of all these other factors, the remaining differences between truants and non-truants could be established. For each outcome variable, separate analyses were carried out for men and women. For discrete outcome variables loglinear analysis was carried out, for continuous outcomes analysis of variance was used.

Information on truancy was available for 10,640 respondents, 85 per cent of the total sample of 12,537 respondents at age 23. However, as the multivariate analyses included a range of variables, there was a higher proportion of cumulative non-response, which varied according to the outcomes evaluated. Analyses were carried out to assess the effects of any non-response bias. It was found that there were no systematic biases which would affect the statistical relationships observed, and that the sample analysed was reasonably representative of the total. For a fuller discussion of non-response and the analyses used to test for its effects, see Hibbett *et al.* (1990).

For comparisons across the stages to be valid, it was clearly important that each stage should be based on the same sample. As the examination and awards system differs between Scotland and the rest of Great Britain, no common measure could be derived for "qualifications" in the fifth stage of the analysis, thus the findings presented here are for England and Wales only. Findings from the separate analysis of continuous variables for Scottish respondents are presented in Hibbett and Fogelman (1988), but the Scottish subsample was too small for loglinear analysis.

RESULTS

Marriage and family formation

In analysing occupational outcomes of truancy, we have suggested that the truants are progressing more rapidly through the major events and transitions of early adult life (Hibbett *et al.*, 1990). Further evidence of this is provided by Tables 2 and 3. The first of these is concerned with legal marital status, the second with current living situation. Findings in this and the following tables are based on loglinear analysis, and report results as relative odds calculated against a selected reference category: 95 per cent confidence intervals are given to indicate statistical significance. As there were found to be no consistent interactions between the truancy variable and other independent variables, main effects only were fitted. In Table 2 we see that, compared with the non-truants, male truants are just over half as likely to be single rather than married, whereas the equivalent figure for women is two thirds. For

TABLE 2

MARITAL STATUS (*LOGLINEAR ANALYSIS*). ODDS RELATIVE TO 'MARRIED'. TRUANTS AT 16 VS. NON-TRUANTS

	Adjusted for				
		Background AND	Educ. Ability AND at 11	Attendance AND at 15	Qualifications at 16+
MEN (2815)					
Single	0.58 (0.47/0.70)	0.66 (0.54/0.81)	0.64 (0.52/0.79)	0.64 (0.52/0.79)	0.64 (0.52/0.80)
Separated/ Divorced/Widowed	2.19 (1.31/3.68)	2.07 (1.21/3.54)	1.99 (1.12/3.52)	2.00 (1.13/3.54)	2.01 (1.12/3.61)
WOMEN (2848)					
Single	0.67 (0.53/0.84)	0.78 (0.61/0.99)	0.79 (0.61/1.02)	0.82 (0.63/1.08)	0.78 (0.58/1.06)
Separated/ Divorced/Widowed	3.50 (2.43/5.05)	3.57 (2.42/5.26)	3.44 (2.34/5.08)	3.33 (2.25/4.93)	3.40 (2.28/5.07)

95 per cent confidence intervals in brackets

men, it is only adjusting for social background which reduces the contrast, and that only marginally. Although the same is true for women, the reduction is sufficient to bring the difference to the borderline of statistical significance. The more dramatic figures in this table relate to the odds of being separated, divorced or widowed (and the last of these is of course very rare at this age). Although overall the number whose marriages have broken up by 23 is not large, the differences are highly significant, and suggest that male truants are about twice as likely to have experienced a broken marriage by the age of 23, and female truants more than three times as likely.

As would be expected, the findings for the odds of not being currently in a partnership ("lone" in Table 3) are similar to those of being single, since they are overwhelmingly the same group. However, some of the single (and some of the separated or divorced) were cohabiting at the time of the 23-year interview, and here there is a marked sex difference in the relationship with truancy. Male truants are slightly less likely to be cohabiting, but the difference is not statistically significant. On the other hand, female truants are much more likely to be cohabiting than are the non-truants.

TABLE 3

PARTNERSHIP STATUS (*LOGLINEAR ANALYSIS*). ODDS RELATIVE TO 'MARRIED'. TRUANTS AT 16 VS. NON-TRUANTS

	Adjusted for				
		Background AND	Educ. Ability AND at 11	Attendance AND at 15	Qualifications at 16+
MEN (2815)					
Lone	0.60 (0.49/0.73)	0.68 (0.56/0.84)	0.66 (0.53/0.81)	0.66 (0.53/0.81)	0.66 (0.53/0.81)
Cohabiting	0.82 (0.54/1.25)	0.96 (0.62/1.47)	0.86 (0.55/1.36)	0.87 (0.55/1.38)	0.87 (0.55/1.38)
WOMEN (2848)					
Lone	0.77 (0.61/0.96)	0.88 (0.70/1.12)	0.89 (0.69/1.14)	0.87 (0.68/1.13)	0.81 (0.62/1.07)
Cohabiting	1.62 (1.14/2.28)	1.86 (1.29/2.70)	1.96 (1.36/2.83)	1.90 (1.30/2.80)	1.79 (1.21/2.66)

95 per cent confidence intervals in brackets

Further insight is provided in Table 4, which summarises analyses of current family type. Although we have seen fairly substantial differences in the likelihood of being married at 23, truants and non-truants differ only slightly in the likelihood of being married without children. There is a much greater contrast in the odds of being married with children and single with children (this last group includes the separated, divorced or widowed and those cohabiting). Female truants are particularly more likely to be in this last situation — as much as six or seven times more likely than non-truants. A small proportion of these differences is explained by social background, but there is no additional contribution from other factors in the analyses.

Looking further at the children of the cohort members, in terms of the cohort member's age when their first child was born, and the number of children they have had by the age of 23, it was found that for those men with children, truants are only about three months younger on average at the time of the birth of their first child — a difference which is not statistically significant and is further reduced when adjusted for other factors (table not presented). Among women the initial difference is larger, about 10 months. Allowing for social background and ability at 11 reduces this to about six months, and allowing additionally for attendance and examination attainment brings it to a level of about four months, which is no longer statistically significant.

CALCUTTA CENTRAL LIBRARY

TABLE 4

Family Type (*Loglinear Analysis*). Odds Relative to 'Single (including Separated/Divorced/Widowed) No Children'. Truants at 16 vs. Non-truants

	Adjusted for				
		Background AND	Educ. Ability AND at 11	Attendance AND at 15	Qualifications at 16+
MEN (2815)					
Single with Children	2.65 (1.46/4.80)	2.07 (1.12/3.82)	2.22 (1.25/3.95)	2.76 (1.62/4.70)	2.83 (1.66/4.83)
Married with Children	2.64 (2.07/3.39)	2.08 (1.61/2.69)	2.41 (1.84/3.14)	2.53 (1.94/3.29)	2.47 (1.88/3.23)
Married No Children	1.21 (0.95/1.54)	1.15 (0.90/1.47)	1.16 (0.90/1.48)	1.15 (0.90/1.47)	1.14 (0.89/1.45)
WOMEN (2848)					
Single with Children	7.55 (5.21/10.93)	5.67 (3.86/8.35)	6.10 (4.00/9.28)	6.06 (3.95/9.28)	6.06 (3.93/9.34)
Married with Children	3.09 (2.38/4.02)	2.27 (1.73/2.97)	2.28 (1.70/3.07)	2.28 (1.70/3.06)	2.28 (1.65/3.14)
Married No Children	0.74 (0.53/1.04)	0.73 (0.52/1.03)	0.74 (0.52/1.04)	0.74 (0.52/1.04)	0.74 (0.52/1.05)

95 per cent confidence intervals in brackets

A similar sex difference is also apparent in relation to the number of children in the family (Table 5). Male truants were over three times as likely and female truants over six times as likely than non-truants to have two or more children by the age of 23. Adjusting for the influence of other factors reduces these odds, but they are still 2.7 and 3.56 for male and female truants respectively. Social background accounts for some of the initial contrast between truants and non-truants for men, with the other remaining factors being of little influence, whereas for women prior educational attainment explains a relatively larger proportion of the initial difference.

TABLE 5

Number of Children (*Loglinear Analysis*). Odds Relative to 'No Children'. Truants at 16 vs. Non-truants

	Adjusted for				
		Background AND	Educ. Ability AND at 11	Attendance AND at 15	Qualifications at 16+
MEN (2815)					
One child	2.42 (1.86/3.15)	1.96 (1.50/2.58)	2.05 (1.54/2.72)	2.05 (1.54/2.73)	2.02 (1.50/2.72)
Two or more children	3.66 (2.40/5.59)	2.68 (1.72/4.15)	2.81 (1.84/4.31)	2.81 (1.82/4.35)	2.70 (1.67/4.38)
WOMEN (2848)					
One child	3.35 (2.60/4.31)	2.57 (1.97/3.34)	2.38 (1.80/3.15)	2.38 (1.80/3.15)	2.36 (1.78/3.13)
Two or more children	6.04 (4.58/7.97)	4.03 (3.00/5.39)	3.59 (2.60/4.97)	3.70 (2.68/5.11)	3.56 (2.47/5.12)

95 per cent confidence intervals in brackets

TABLE 6

SMOKING (*LOGLINEAR ANALYSIS*). ODDS RELATIVE TO 'NON-SMOKING'. TRUANTS AT 16 VS. NON-TRUANTS

	Adjusted for				
		Background AND	Educ. Ability AND at 11	Attendance AND at 15	Qualifications at 16+
MEN (2815)					
Up to 10 cigarettes per day	2.00 (1.37/2.93)	2.18 (1.47/3.23)	2.11 (1.43/3.13)	2.02 (1.38/2.98)	2.05 (1.39/3.02)
10 - 30 cigarettes per day	3.42 (2.77/4.22)	3.11 (2.51/3.86)	3.12 (2.51/3.87)	3.19 (2.57/3.97)	3.28 (2.60/4.14)
Over 30 cigarettes per day	5.79 (3.83/8.76)	4.89 (3.19/7.54)	4.93 (3.20/7.58)	4.94 (3.25/7.54)	5.05 (3.28/7.85)
WOMEN (2848)					
Up to 10 cigarettes per day	1.82 (1.25/2.64)	2.06 (1.41/3.03)	2.16 (1.46/3.18)	2.17 (1.48/3.20)	1.83 (1.22/2.73)
10 - 30 cigarettes per day	4.44 (3.50/5.63)	3.85 (3.02/4.92)	4.05 (3.13/5.22)	4.14 (3.21/5.34)	4.34 (3.39/5.61)
Over 30 cigarettes per day	8.89 (5.26/15.00)	7.65 (4.39/13.33)	8.72 (5.05/15.09)	9.47 (5.54/16.12)	10.51 (6.17/17.81)

95 per cent confidence intervals in brackets

Health-related behaviour

As can be seen in Table 6, truants of both sexes are considerably more prone to smoking and to smoking heavily (over 30 cigarettes per day). They are correspondingly much less likely to be non-smokers. This appears to be more pronounced for female truants. For male truants the initial odds of smoking 10-30 cigarettes, or over 30 cigarettes per day, are reduced slightly by allowing for other factors, but for women there is no consistent reduction in the initial differences by accounting for other influences. Indeed, the likelihood of smoking more than 30 cigarettes a day is, with 10.51, higher after adjusting for other factors than the initial one of 8.89, indicating that among women of similar background, ability, attendance and qualifications, truants have an increased likelihood of smoking heavily. By contrast truants differ hardly at all from non-truants in their drinking habits (table not presented). The confidence intervals indicate that the only noteworthy contrast is that female truants are more

TABLE 7

DEPRESSION (MALAISE-SCORE CATEGORISED) (*LOGLINEAR ANALYSIS*). ODDS RELATIVE TO 'NORMAL'. TRUANTS AT 16 VS. NON-TRUANTS

	Adjusted for				
		Background AND	Educ. Ability AND at 11	Attendance AND at 15	Qualifications at 16+
MEN (2813)					
Depressed	2.64 (1.75/3.98)	2.26 (1.48/3.47)	2.24 (1.37/3.68)	2.38 (1.47/3.85)	2.40 (1.48/3.90)
WOMEN (2842)					
Depressed	2.75 (2.08/3.63)	2.16 (1.61/2.90)	2.26 (1.68/3.06)	2.23 (1.63/3.06)	2.22 (1.61/3.07)

95 per cent confidence intervals in brackets

likely to report that they drink less often than once or twice a week or never, but even these differences are no longer significant once social background is taken into account.

Finally, the analysis reported in Table 7 is of scores on the "Malaise Inventory" (Rutter, 1976), a screening instrument designed to detect tendency to depression. Differences are substantial for both sexes, with truants almost three times as likely to obtain scores above the cut-off indicating depression. The effect on this of taking social background into account is small, and of other factors trivial.

DISCUSSION

Before discussing our findings, it is important to note that, when looking at outcomes, the age of 23 is clearly relatively early in the adult lives of the cohort. By the age of 23 about half the cohort were married, and about one third had had children. Therefore we have not examined a static or completed picture. It should also be noted that progress through the transitions of early adulthood varies according to social class and sex. Young people from working class backgrounds tend to leave the parental home, marry and have children earlier, as do young women compared to men. And women are more likely to leave the labour force, if only temporarily, when they have young children. These general patterns in the transitions of early adulthood (Kiernan, 1986) may explain some of the differences observed between truants and non-truants, and male and female respondents, and also emphasise the need for taking these factors into account by using a multivariate design.

We found that truants at 16 differed consistently and significantly from non-truants in terms of family formation and health-related outcomes, up to the age of 23. They tend to marry and have children at an earlier age and by 23 they have more children. They are by this age considerably more likely to be separated or divorced, often with children in their care. On the whole, truants appear to put a stronger emphasis on having children and getting married, which may be in compensation for previously observed occupational disadvantages (Hibbett *et al.*, 1990). The results of the preliminary bivariate analyses suggest that such differences are even more marked for the relatively small number who had been identified as truanting at both 11 and 16.

Truants were found to be much more prone to heavy smoking, and this contrast was relatively unaffected by controlling for other factors. This study therefore confirms the recent findings of Charlton and Blair (1989), establishing an association between smoking and truancy. There was little difference between truants and non-truants regarding frequency of drinking. This is in line with previous research in this country (Farrington, 1980), but appears to contradict the findings of the American study by Robins and Ratcliffe (1980), who found truants to have a higher likelihood of subsequent alcohol problems.

It was found that, generally, social background and prior educational attainment contributed only a small part to explaining the contrasts between truants and non-truants, which still left a substantial and significant difference between the two groups. The effect on the contrast between truants and non-truants of allowing for attendance rate in the autumn term of the final compulsory year is also on the whole insignificant. Likewise, the remaining differences between truants and non-truants are not explained by accounting for the effects of educational qualifications. This means that there are consistent and significant differences in the outcomes examined between truants and non-truants of similar background, ability, attendance and educational attainment.

An explanation might be that the difficulties encountered by the former truants are the result of their more accelerated rate of progress through the major events and transitions of early adulthood. They were married at an earlier age, were younger when they had their first child and had more children by the age of 23. In these respects they are showing a more extreme version of the pattern of other young people of similar background. Even the higher rates of marital breakdown of the truants may be compatible with their having reached this stage at an earlier age; their peers may experience similar problems as they grow older.

However, when male former truants are twice as likely to have experienced marital breakdown, and females three times as likely, it does not seem probable that these are difficulties only temporarily greater than those of their peers. On the contrary, it seems more reasonable to interpret this as one symptom of their generally greater problems in adult life. A further important sign of this, in addition to their employment difficulties previously observed, are their high scores on the "Malaise" inventory. Although not synonymous with clinical depression, scores above the conventional cut-off do indicate a greater likelihood of depression, and were more than twice as frequent for the truants. Interestingly, while this is generally in line with previous research (Tyrer and Tyrer, 1974; Rodgers, 1990), we found male and female truants to be similar in this respect.

Conclusions

Our findings suggest that there are significant and consistent differences between truants and non-truants, regarding family formation and health-related behaviour. The much higher incidence of subsequent marital breakdown and depressive symptoms for truants suggest that truancy is part of more comprehensive and persistent psychological problems, rather than merely a move away from the world of school to the world of work. In our previous paper (Hibbett *et al.*, 1990) we noted that truancy is associated with serious occupational disadvantages by age 23; the present analyses extend and consolidate these observations, indicating that truancy is a predictor of multiple problems in early adulthood.

ACKNOWLEDGMENTS. — We are grateful to the ESRC who funded this study, and to our colleagues in the Social Statistics Research Unit, City University, for their advice and support.

Correspondence and requests for reprints should be addressed to Dr Angelika Hibbett, Social Science Branch 1, Employment Department, Steel House, Tothill Street, London SW1H 9NF.

REFERENCES

Charlton, A., and Blair, B. (1989). Absence from school related to children's and parental smoking habits. *Br. med. J.*, 298, 90-92.

Cherry, N. (1976). Persistent job changing — is it a problem? *J. occup. Psychol.*, 49, 203-221.

Farrington, D. (1980). Truancy, delinquency, the home and the school. In Hersov, L., and Berg, I. (Eds.), *Out of School*. Chichester: John Wiley and Sons.

Fogelman, K. (1978). School attendance, attainment and behaviour. *Br. J. educ. Psychol* , 48, 148-158.

Fogelman, K., and Richardson, K. (1974). School attendance: some results from the National Child Development Study. In Turner, B. (Ed.), *Truancy*. London: Ward Lock Educational.

Gray, G., Smith, A., and Rutter, M. (1980). School attendance and the first year of employment. In Hersov, L., and Berg, I. (Eds.), *Out of School*. Chichester: John Wiley and Sons.

Haigh, G. (1976). *The Reluctant Adolescent*. Aldershot: Temple Smith.

Hibbett, A. (1987). *Early Adult Outcomes of Truancy*. NCDS User Support Group Working Paper 24, Social Statistics Research Unit, City University.

Hibbett, A., and Fogelman, K. (1988). *Early Adult Outcomes of Truancy II: The Effects of Truancy after Allowing for Other Factors*. NCDS User Support Group Working Paper 30, Social Statistics Research Unit, City University.

Hibbett, A., Fogelman, K., and Manor, O. (1990). Occupational outcomes of truancy. *Br J. educ Psychol.*, 60, 23-36.

Kiernan, K. (1986). *Transitions in Young Adulthood*. NCDS User Support Group Working Paper 16, Social Statistics Research Unit, City University.

Robins, L. N., and Ratcliffe, K. S. (1980). The long-term outcome of truancy. In Hersov, L. and Berg, I. (Eds.), *Out of School*. Chichester: John Wiley and Sons.

Rodgers, B. (1990). Behaviour and personality in childhood as predictors of adult psychiatric disorder. *J. Child Psychol. Psychiat.*, 31, 393-414.

Rutter, M., Tizard, J., Yule, W., and Graham, P. (1976). Isle of Wight studies: 1964-1974. *Psychol. Med.*, 6, 313-332.

Shepherd, P. (1985). *The National Child Development Study: An Introduction to the Origins of the Study and the Methods of Data Collection*. NCDS User Support Group Working Paper 1, Social Statistics Research Unit, City University.

Tyerman, M. J. (1968). *Truancy*. London: University of London Press.

Tyrer, P., and Tyrer, S. (1974). School refusal, truancy and adult neurotic illness. *Psychol. Med.*, 4, 416-421.

Watters, J. (1989). School refusal. *Br med. J.*, 298, 66-67.

(*Manuscript received* 15th *September*, 1989)

Br. J. educ. Psychol., 60, 180-191, 1990

THE EXAMINATION ATTAINMENTS OF SECONDARY SCHOOL PUPILS WHO ABUSE SOLVENTS

BY OLIVER CHADWICK
(Department of Psychology, Springfield Hospital, Tooting, London)
WILLIAM YULE
(Department of Psychology, Institute of Psychiatry, London)
AND ROSS ANDERSON
(Department of Clinical Epidemiology and Social Medicine, St George's Hospital Medical School, London)

SUMMARY. A sample of 105 adolescents who had abused solvents and an equal number of controls matched for school, school year and sex were identified by means of a two-stage screening procedure involving (a) administration of a self-completed questionnaire to approximately 5,000 third, fourth and fifth year pupils in 16 secondary schools, and (b) individually conducted interviews. These two groups were followed up at the end of their final year of compulsory education when details of their fifth year public examination entries and results were obtained from their schools. The examination attainments of the solvent abusers were generally lower than those of the controls. They were more likely to leave school at Easter of their fifth year and a much higher proportion of them were not entered for any examinations. Those who were entered for exams were less likely to pass at O level and, regardless of the type of exam, much less likely to obtain higher examination grades. Poor examination attainments among the solvent abusers were associated with adverse domestic circumstances, frequent self-reported behaviour problems and low scores on educational tests administered at the time of transfer from primary to secondary school. After controlling for these factors, the difference between groups in examination attainments was no longer statistically significant.

INTRODUCTION

CONCERN has frequently been expressed about the risks of solvent abuse. Instances of sudden death among youngsters who have been sniffing have been well documented (Bass, 1970; Anderson *et al.*, 1985) and a number of reports have drawn attention to the health risks of the practice (Watson, 1986).

Less attention has been paid to the impact of solvent abuse on educational progress. Only a few studies have examined the relationship between solvent abuse and scholastic attainment and these have mostly been concerned with delinquent samples. Barker and Adams (1963), for example, studied boys at a correctional school in Colorado and found that those who had been committed for glue-sniffing were more educationally retarded than non-sniffing delinquents; both groups were behind in their grade placements, but the glue-sniffers were on average two years further behind than the controls. Similarly, Sterling (1964) in his comparison between juveniles arrested for glue-sniffing or drinking found that the sniffers tended to have lower grade placements than the drinkers.

Very few studies have been conducted with non-delinquent samples and it is therefore unclear whether the reported relationship between solvent abuse and poor scholastic performance holds true for the general population. Nurcombe *et al.* (1970) studied boys at an aboriginal primary school in Australia and noted a trend for those who were said to have been involved in petrol sniffing to be rated as "slow learners" more often than non-sniffing controls. However, both the ratings of learning disability and the judgments about involvement in petrol sniffing were made by boys' teachers and the possibility of unintentional bias cannot be ruled out. Mahmood (1983) found that the mean reading age of a group of solvent abusers referred to a special clinic in Glasgow was approximately four years behind that of a small control group of local secondary school pupils. However, the overall pattern of results in this study is puzzling as the solvent abusers obtained significantly *higher* scores than the controls on tests of arithmetic and paired associate learning.

These few studies do not provide a firm basis for the belief that solvent abuse is associated with poor educational attainment. In view of the widespread nature of the problem, there is a need to confirm this association in a larger and more representative sample of secondary school pupils.

Assuming that there is an association between solvent abuse and poor scholastic attainment, a further question concerns the nature of the hypothesised relationship. Several possible explanations may be suggested. In view of the evidence linking solvent abuse with neurological impairment (Ron, 1986; Lolin, 1989), one possibility is that the association reflects a direct causal effect of solvent abuse on cognitive ability. Alternatively, the poor educational attainments of solvent abusers may be a reflection of background social disadvantage, concomitant behaviour problems or their level of scholastic ability prior to taking up the practice. A large number of studies have shown that poor scholastic performance is more common among children with behavioural problems and among those coming from homes characterised by a variety of social disadvantages (e.g., Rutter *et al.*, 1970). In seeking to understand the relationship between solvent abuse and poor scholastic attainment, the role of each of these factors has to be considered.

The aim of the present study was, therefore, to examine in a representative sample of pupils attending ordinary secondary schools whether there is an association between solvent abuse and poor educational attainment, and, if so, to investigate whether this association is due specifically to solvent abuse.

METHOD

Children who had abused solvents were provisionally identified by means of a self-completed screening questionnaire that was administered to the third, fourth and fifth year pupils (aged 13-16 years) of 16 local education authority secondary schools in three geographically adjacent areas: one division of the Inner London Education Authority (five schools) and two outer London boroughs (11 schools). At one school the questionnaire was administered at the end of the school year 1983/4 but at the other 15 it was administered during the school year 1984/5. At schools where the questionnaire survey was conducted during the summer term, fifth year pupils were excluded.

At each school, a letter was sent to the parents of each child in the target population explaining the study and giving them the opportunity to withdraw their child. Pupils whose parents refused permission were excluded form the study, as were those pupils who themselves declined to participate.

The questionnaire was administered to the pupils in their normal classrooms and was introduced, distributed and collected by a member of the research team. It took about 15-25 minutes to complete and consisted of 31 items, three of which were concerned with solvent abuse. In introducing the questionnaire, the confidentiality of the pupils' responses was stressed, and to emphasise the point, their teachers were openly encouraged to feel free to leave the classroom as soon as the class had settled. The pupils were requested to record their date of birth and school class on the questionnaire so that those selected for further assessment and follow-up could later be identified by means of the school register.

A total of 7,485 third, fourth and fifth year pupils from the 16 participating schools were eligible for inclusion in the questionnaire survey. At the request of their parents 825 of them (11.0 per cent) were excluded, and 1,180 (15.8 per cent) were recorded as absent on the morning or afternoon the questionnaire was administered. Completed questionnaires were obtained from a total of 5,014 pupils.

Individual assessments

At each school, a sample of those who acknowledged that they had deliberately abused solvents to the point of intoxication were selected for individual interview. At schools where the questionnaire yield was high, those who indicated that they had abused solvents most frequently were examined first in order to ensure that as many frequent abusers as possible were assessed in the time available.

For each case, a control who denied sniffing on the questionnaire was also selected for interview. The controls were selected at random and were individually matched with the cases for sex, school year and school.

The aim of the interview was to provide a check on the validity of the questionnaire responses to the items on solvent abuse, to obtain detailed information on the frequency and pattern of solvent abuse, and to collect information on background social and medical factors that might be expected to affect educational progress. Although the structure of the interview followed a standard format, the interviewers were encouraged to supplement the basic schedule in any way they judged likely to set the interviewee at ease, promote confidence and thereby increase the likelihood of obtaining valid and truthful information.

One hundred and thirty three positive questionnaire respondents were individually interviewed and solvent abuse was confirmed at interview by 106 (79.7 per cent) of them. One hundred and six negative questionnaire respondents were also interviewed and all of them confirmed that they had not abused solvents. The two groups therefore consisted of 106 interview-confirmed cases and a matched sample of 106 interview-confirmed controls.

The individual assessments also included an extensive battery of neuropsychological tests and a short breath test, the aim of which was to identify pupils who might have been under the influence of solvents (or alcohol) at the time of psychological testing. Detailed results from these assessments are presented elsewhere (Chadwick *et al.*, 1989a). For the present, it is sufficient to note that seven of the cases and one of the controls were positive for solvents on the breath test. The control (and the case with whom he was matched) were excluded from their respective samples, thereby reducing the sample size to 105 matched pairs.

Fifth year public examination results

At the end of the school years 1984/5, 1985/6 and 1986/7, the headteachers of each of the participating schools were requested to provide details and results of all public examinations taken by the pupils who had been individually assessed. A standard form was used for this purpose with space to record each subject entered for, at what level (CSE, GCE O-level, or '16+'), on what date and with what result. Where pupils were entered for more than one level of examination in the same subject (e.g., CSE and O-level Maths), only the better result was counted. Information was also collected on whether the pupil left school at Easter or in the summer of their fifth year or whether he or she stayed on into the sixth form. Although the headteachers were aware that the study involved a comparison between pupils who had abused solvents and those who had not, at no point was information disclosed about the group membership of particular pupils.

All headteachers provided public examination data, and information was obtained on 102 cases and 103 controls. The remaining three cases and two controls had transferred to other schools before Easter of their fifth year. No attempt was made to collect examination results from their new schools because permission to do so had not been obtained at the outset.

Results on three types of examination were examined: Certificate of Secondary Education (CSE), General Certificate of Education at Ordinary Level (GCE O-level) and a precursor (16+) of the General Certificate of Secondary Education (GCSE). GCE O-levels are marked on a five-point scale ranging from A to E and are usually taken by the more academically able pupils: at least five passes at grade C or above are normally required for admission to higher education and entry into certain careers. CSE is a newer type of examination which aims to assess a wider range of scholastic abilities. It is marked on a 1 to 5 scale. There is an accepted scale of equivalence between O-level and CSE exam marks (e.g. grade 1 at CSE is usually considered equivalent to grade C at O-level). The GCSE examination was officially introduced shortly after this study was completed with the aim of replacing the GCE and CSE curricula and exams with a new curriculum and a single scale of scholastic attainment. At the time the study was carried out, some schools offered their pupils the option of taking a precursor of the GCSE (16+ examination) in some subjects. These 16+ examinations were graded by the examiners in terms of both CSE and O-level grade equivalents. Usually these grade equivalents corresponded closely, but where discrepancies occurred, the O-level grade equivalent was used in the present study.

Educational test results at time of secondary school entry

In order to obtain data on educational test performance prior to solvent abuse, the schools were asked to provide details of the results on any standardised educational tests that had been taken by the cases and controls at the time of transfer from primary to secondary school. With the cases, the date when the test was taken was compared with the age when the child had started to abuse solvents (as determined at interview) to ensure that the test had indeed been taken prior to the first episode of solvent abuse.

Valid data came from (a) tests of reading ability (Neale Analysis of Reading Ability, Schonell Word Reading Test, Holborn Reading Test, London Reading Test or Burt Word Reading Test) or (b) verbal reasoning tests (NFER VR Tests, Cognitive Abilities Test (VQ) and the Moray House VR Test). On most of the tests, the data came in the form of an age standardised quotient with a mean of 100 and a standard deviation of 15, but on the Neale, Schonell and Holborn reading tests, the scores came in the form of a reading age (in years and months). In order to convert these two types of score to the same metric, the reading ages were transformed into equivalent scores on the London Reading Test Form A (1980) using the empirical data collected to validate the test, and converted into standardised quotients.

Behavioural and emotional problems

Information on behavioural and emotional problems was obtained on the self-completed screening questionnaire by means of a 15-item checklist (Chadwick, 1988), modelled on Rutter's (1967) Child Behaviour Questionnaire (B2), covering problems with sleeping, appetite, stomach aches, headaches, nervousness, difficulties concentrating, worrying, being bullied, involvement in fights, stealing, destructive behaviour, truancy, misery, loneliness and worries about health.

RESULTS

Frequency and pattern of solvent abuse

Before presenting the examination results, it is important to mention the frequency and pattern of solvent abuse reported by the cases at interview. By design, the proportion of frequent solvent abusers was somewhat higher among the interviewed cases than among the total sample identified by means of the questionnaire (Chadwick *et al.*, 1989a). Even so, the pattern of solvent abuse among the former group could best be described as "experimental". Approximately a third of those who had abused solvents had done so only once or twice, and nearly two-thirds of them had sniffed fewer than ten times. At the other end of the frequency distribution, approximately one in five had sniffed more than 30 times. Typewriter correction fluids were the most popularly abused products and had been tried by approximately three-quarters of the sample. Various glues and butane gas had each been abused by between a quarter and a third of them. Other substances, such as dry cleaning fluids, amyl nitrite, petrol, paints and aerosols had been abused by a smaller proportion of the sample.

Time of school-leaving

As shown in Table 1, the solvent abusers were significantly more likely to leave school at an earlier stage than the controls ($\chi^2 = 12.51$, 2 df: $P<0.01$).

TABLE 1

TIME OF SCHOOL LEAVING

	Control (N = 103)	Solvent Abusers (N = 102)
Left Easter of 5th year	4 (4.0%)	19 (19.2%)
Left Summer of 5th year	68 (68.0%)	61 (61.6%)
Went into 6th form or equivalent	26 (26.0%)	16 (16.2%)
Not known	5	6

Examination entries

The solvent abusers were entered for significantly fewer examinations (mean = 4.0; SD = 3.3) than the controls (mean = 6.0, SD = 2.3; t = 4.96, 203 df $P<0.0001$). However, a large part of this difference could be explained in terms of differences between the two groups in the proportion of children who were not entered for any exams. Whereas 30.4 per cent of the cases were entered for no exams, this was true of only 3.9 per cent of the controls, and when these children were omitted from the comparison, the difference in mean number of examination entries (5.8, SD = 2.3 vs. 6.2, SD = 1.9) was no longer statistically significant.

There was a statistically significant difference between the two groups in terms of the levels of examinations entered for ($\chi^2 = 6.88$, 2 df: $P<0.05$). The examination entries of the solvent abusers were less likely to be for O-levels than the controls (16.0 vs. 22.5 per cent) and more likely to be for CSEs (72.1 vs. 67.7 per cent) and 16+ exams. (11.9 per cent vs. 8.8 per cent) instead (Chadwick, 1988).

Examination absenteeism

The solvent abusers were somewhat more likely to be absent for examinations than the controls. Among those who were entered for any exams, 16 (22.5 per cent) of the 71 solvent abusers and 9 (9.1 per cent) of the 99 controls were absent for at least one exam [$\chi^2 = 3.18$, 1 df, NS]. Most of these absentees were absent for only one or two exams, and overall, the solvent abusers were absent for 24 (5.8 per cent) of the 413 exams for which they were collectively entered whereas the controls were absent for 19 (3.1 per cent) of their 619 exam entries.

Examination results

As shown in Table 2, the overall examination performance of the controls was markedly better than that of the solvent abusers ($\chi^2 = 25.10$, 4 df, $P<0.0001$). The main difference between the two groups appeared to lie at the extremes of the scale.

For the purposes of statistical analysis, there are advantages to the use of a continuous measure of examination performance. The results were therefore reanalysed using the "weighted examination score" proposed by Byford and Mortimore (1980) which takes account of the grades obtained on each exam as well as the number of exams passed. Points

TABLE 2

OVERALL CLASSIFICATION OF EXAMINATION ATTAINMENTS

	Controls (N = 100)	Solvent Abusers (N = 99)
No graded pass	9 (9.0%)	34 (34.3%)
CSE grade 4 or 5	10 (10.0%)	10 (10.1%)
CSE grades 2 - 3 or O-level grades D - E	35 (35.0%)	29 (29.3%)
1 - 4 O-levels grade A - C or CSEs grade 1	35 (35.0%)	25 (25.2%)
5+ O-levels grade A - C or CSEs grade 1	11 (11.0%)	1 (1.0%)

NOTE: Three children in each group who were absent for all exams entered for are excluded from this Table.

are allocated for each examination grade as shown in Figure 1, and the weighted score is obtained by summing the scores on all exams taken.

FIGURE 1

POINTS ALLOCATED IN CALCULATING WEIGHTED EXAMINATION SCORES

Examination	Grade							
O-level	A	B	C	D	E	—	—	U
CSE	—	—	1	2	3	4	5	U
Points	7	6	5	4	3	2	1	0

The mean weighted examination score of the solvent abusers (10.4, SD = 10.3) was significantly lower than that of the controls (18.4, SD = 12.9, t = 4.9, df = 189, P<0.0001).

In the above comparisons, children who were not entered for any examinations were regarded as having achieved no graded passes. In order to examine whether the large differences in examination results between groups could be explained in terms of the greater number of examination entries among the controls, the comparisons were repeated after taking into account the number of examinations taken by each group.

Table 3 shows the aggregate number of examination passes expressed in relation to the aggregate number of examinations taken by each group. The overall pass rate among the solvent abusers was significantly lower than the pass rate among the controls ($\chi^2 = 5.63$, 1 df, $P<0.025$) and this difference was almost entirely attributable to their significantly lower O-level pass rate ($\chi^2 = 15.03$, 1 df: $P<0.001$). The difference between groups in CSE and 16+ pass rates were negligible.

In order to take into account the effect of the number of examination entries on the weighted examination scores, the score of each child was divided by the number of examinations taken to produce a mean weighted score for each child. The mean of the solvent abusers' mean scores (2.41, SD = 1.08) was lower than that of the controls (2.78, SD = 1.43), but this difference fell just short of statistical significance (t = 1.88, df = 161, P<0.062).

TABLE 3

Proportion of Exam Entries Passed

	Controls			Solvent Abusers		
	Number of Exams			Number of Exams		
	Passed	Taken	%	Passed	Taken	%
CSEs	364	402	(90.5%)	248	279	(88.9%)
O-levels	132	138	(95.6%)	48	64	(75.0%)
16+	56	60	(93.3%)	44	46	(95.6%)
Total	552	600	(92.0%)	340	389	(87.4%)

Other factors associated with both solvent abuse and examination performance

(a) *Social disadvantage*

The domestic circumstances of the solvent abusers were generally less favourable than those of the controls. As shown in Table 4, they were slightly less likely to be living with both their natural parents, slightly more likely to come from families living in rented accommodation and significantly more likely to be members of large sibships than the controls ($\chi^2 =$ 4.52, 1 df, P<0.05). In addition, they were slightly less likely to come from homes in which the main breadwinner had a non-manual occupation, and significantly more likely to come from homes in which neither parent was employed ($\chi^2 = 6.61$, 1 df: P<0.01).

Adverse domestic circumstances were in turn associated with poor examination results among the solvent abusers (see Table 5). Children from large sibships (t = 3.02, 97 df, P<0.01), those whose families lived in rented accommodation (t = 2.06, 94 df, P<0.05), and those from families in which the main breadwinner had a manual occupation (t = 2.28, 75 df P<0.05) obtained significantly lower weighted examination scores than those coming from homes without these features.

(b) *Behavioural and emotional problems*

As shown in Table 6, most of the items on the behavioural and emotional problems checklist were acknowledged as being a frequent problem significantly more often by the

TABLE 4

Background Domestic Factors

	Controls		Solvent Abusers		P<
Not living with both 'natural' parents (Q)	35/103	(34.0%)	41/101	(40.6%)	NS
3 or more sibs (I/V)	27/103	(26.2%)	41/102	(40.2%)	0.05
Accommodation owned by family (Q)	62/103	(60.2%)	47/100	(47.0%)	NS
Occupational status of main breadwinner (I/V)					NS
I, II, III Non-manual	36/96	(37.5%)	21/80	(26.2%)	
III Manual	40/96	(41.7%)	44/80	(55.0%)	
IV, V	20/96	(20.8%)	15/80	(18.7%)	
Employment status of main breadwinner (I/V)					
Unemployed	3/99	(3.0%)	12/92	(13.0%)	0.01

(Q) : Questionnaire data
(I/V) : Interview data

solvent abusers than by the controls. In addition, at interview the solvent abusers were significantly more likely to express antipathy towards their teachers(see Table 7).

The associations between the presence of each of these 16 problems (15 from the checklist and one from the interview) and the pupils' weighted examination scores were examined among the solvent abusers. The only items that were consistently associated with low examination scores were those relating to anti-social problems: fighting, stealing, destructive behaviour, truancy and antipathy towards teachers. Cases who acknowledged these problems obtained significantly lower weighted examination scores (mean = 7.6, SD = 10.1) than the rest of the cases (mean = 12.6, SD = 10.1, t = 2.48, df = 97, P<0.015).

TABLE 5

Background Domestic Factors and Examination Attainments in Solvent Abusers

	N	Weighted Exam Score Mean	(SD)	P<
Living with both 'natural' parents (Q)				
Yes	60	10.5	(10.4)	NS
No	41	10.1	(10.4)	
Number of sibs (I/V)				
<3	61	12.6	(10.8)	0.01
3+	41	6.5	(8.4)	
Housing tenure (Q)				
Owned	47	12.5	(11.7)	0.05
Rented	51	8.2	(8.7)	
Occupational status of main breadwinner (I/V)				
I - III NM	21	14.1	(11.1)	0.05
III Manual	44	8.2	(9.1)	
IV, V	15	8.9	(10.3)	
Employment status of main breadwinner (I/V)				
Employed	80	9.9	(10.1)	NS
Unemployed	12	7.8	(9.2)	

(Q) : Questionnaire data
(I/V) : Interview data

(c) *Educational test results antecedent to solvent abuse*

Perhaps surprisingly, there was virtually no difference between the solvent abusers and their controls in mean antecedent educational test quotients at the time of transfer from primary to secondary school (mean score of the cases = 103.2, SD = 14.1, mean score of the controls = 103.8, SD = 12.6). The data were re-examined to investigate whether, in spite of the similarity in mean scores, there were differences between groups in the proportion of children with low test scores. Test quotients of less than 85 were obtained by approximately twice as many solvent abusers (12.3 per cent) as controls (6.2 per cent), but this difference was not statistically significant ($\chi^2 = 1.83$, 1 df, NS).

There was, however, a significant correlation (r = 0.48) between the pupils' weighted fifth year examination scores and antecedent educational test quotients (measured approximately four years earlier).

TABLE 6

BEHAVIOURAL AND EMOTIONAL PROBLEMS

	Controls	Solvent Abusers	P<
Sleeping difficulties	13/102 (12.7%)	16/101 (15.8%)	NS
Not eating	13/103 (12.6%)	27/102 (26.5%)	0.01
Stomach aches	9/103 (8.7%)	25/100 (25.0%)	0.01
Headaches	21/102 (20.6%)	27/ 99 (27.3%)	NS
Tense/nervous	10/102 (9.8%)	15/ 98 (15.3%)	NS
Poor concentration	17/103 (16.5%)	43/101 (42.6%)	0.001
Worrying	14/102 (13.7%)	27/102 (26.4%)	0.05
Teased/bullied	4/103 (3.9%)	5/100 (5.0%)	NS
Involved in fights	1/103 (1.0%)	19/101 (18.8%)	0.001
Stealing	1/102 (1.0%)	15/100 (15.0%)	0.001
Destructive	0/102 (0.0%)	8/100 (8.0%)	0.01
Truancy	4/101 (4.0%)	26/102 (25.5%)	0.001
Misery	10/103 (9.7%)	18/101 (17.8%)	NS
Loneliness	6/102 (5.9%)	15/101 (14.8%)	0.05
Worried about health	7/102 (6.9%)	15/102 (14.7%)	NS
	Mean (SD)	Mean (SD)	
Total score	13.7 (7.1)	20.3 (6.8)	0.001

TABLE 7

RELATIONSHIPS WITH TEACHERS

	Controls (N = 103)		Solvent Abusers (N = 102)	
	N	%	N	%
Gets on reasonably well with most teachers	92	(93.9%)	60	(63.2%)
Some problems with most teachers or marked antipathy towards a few	4	(4.1%)	25	(26.3%)
Marked antipathy towards most teachers	2	(2.0%)	10	(10.5%)
Not known	5		7	

$\chi^2 = 27.2$, 2 df: $P < 0.001$

Controlling for differences in social disadvantage, behaviour problems and antecedent educational test results

Taken together, the findings presented above show that in addition to the differences between groups in fifth year examination results, there were also differences in levels of social disadvantage, frequency of self-reported behaviour problems and educational test results prior to solvent abuse that might contribute to, or account for, the poor examination performance of the solvent abusers. In order to examine whether the difference between groups remained after controlling for these potentially confounding variables, the weighted examination scores of the two groups were compared in an analysis of covariance in which these variables were included in the equation as covariates.

Housing tenure (owned or rented), number of siblings (three plus or fewer), and the occupational status of the main family breadwinner (manual or non-manual) were dummy-coded and treated as covariates in the equation to control for differences in background social disadvantage. An overall anti-social problems score ranging from 0 to 14 was constructed by aggregating each pupil's self-rating of the frequency of fighting, stealing, destructive behaviour and truancy (scored on a 0 to 3 scale) and antipathy towards teachers (scored on a 0 to 2 scale). The pupils' antecedent educational test quotients were also included as a fifth covariate. When the two groups were compared after controlling for differences on these five variables, no significant difference between the solvent abusers and their controls was found ($F(1, 110) = 2.92$, $P<0.090$).

DISCUSSION

At the end of their final year of compulsory schooling, the examination attainments of secondary school pupils who had abused solvents were found to be generally lower than those of other randomly selected pupils from the same schools. Taken in conjunction with their roughly normal levels of verbal ability both at the time of secondary school entry and when examined during the course of this study (Chadwick *et al.*, 1989a), their poor examination results would appear to reflect a marked degree of scholastic underachievement.

It might be argued that the sample of solvent abusers studied here was unrepresentative of school age solvent abusers in general because pupils who were absent from school on the day the screening questionnaire was administered were not included. Youngsters who have abused solvents frequently or over prolonged periods of time are probably more likely than others to be absent from school on any given day (cf. Kandel 1975), and as a result, cases of this kind are likely to have been under-represented in the sample. However, as pointed out earlier, pupils who acknowledged frequent solvent abuse on the screening questionnaire were preferentially selected for individual assessment. To some extent, this will have corrected the under-representation of frequent solvent abusers due to school absenteeism (Chadwick *et al.*, 1989b). Moreover, it is important to note that any such under-representation should lead to an overestimate of the examination attainments of the group as a whole. In other words, the examination attainments of the solvent abusers should, according to this argument, be even worse than they were found to be here.

One possible explanation of the findings is that scholastic underachievement represents a long-term effect of neuropsychological impairment caused by solvent abuse. However, as reported elsewhere (Chadwick *et al.*, 1989a), when the two groups were compared on an extensive battery of neuropsychological tests, the few differences that were found (on tests of vocabulary, Verbal IQ and a measure of impulsiveness) could be adequately explained in terms of differences between groups in background social disadvantage.

Alternatively, the poor scholastic performance of the solvent abusers might simply reflect their level of ability prior to solvent abuse. In the present study, the evidence relevant to this possibility was equivocal: although low scores on tests of verbal reasoning and reading ability taken at the time of secondary school entry were twice as common among the solvent abusers as among the controls, the mean test scores of the two groups were closely comparable. However, prior educational test data were not obtained for all pupils and those that were obtained came from a variety of tests administered at different times and presumably in slightly different ways in each school. In view of Seltzer and Oechsli's (1985) finding that adolescent cigarette smokers obtain lower scores than non-smokers on tests of intellectual ability administered in childhood, it is possible that pre-existing weaknesses in verbal reasoning and reading ability among some of the solvent abusers may have contributed to the poor scholastic performance of the group as a whole.

A further possibility is that the poor fifth year examination attainments of the solvent abusers reflect a disaffection toward school values and a disinclination to apply their abilities to school work. The solvent abusers were much more likely to acknowledge both emotional and behavioural problems than the controls, and as in studies of other groups (e.g. Rutter *et al.*, 1979; Maughan *et al.*, 1985), behavioural problems were particularly strongly associated with poor educational attainments. It seems likely that such problems both contributed to, and developed in reaction to, lack of scholastic success. It has been suggested that opposition to authority and adult mores is central to the shared value-system of solvent abuse peer groups (e.g., Merrill, 1985; Watson, 1986). Assimilation into, or involvement in, such peer groups might therefore have the effect of fostering the development of a range of "anti-school" sentiments, values and behaviours. These sentiments would include indifference or hostility to notions of scholastic commitment and achievement.

In keeping with these possibilities, when levels of problem behaviour, background social disadvantage and educational test results at the time of secondary school entry were taken into account, the difference between groups in weighted examination scores was no longer statistically significant. The inference that the relationship between solvent abuse and educational outcome can be adequately explained in terms of these factors is consistent with the conclusions from studies of the impact of marijuana use (Robins *et al.*, 1970; Halikas *et al.*, 1983) and other forms of illicit drug use (Kandel *et al.*, 1986) on educational attainment.

Nevertheless, it is important to emphasise the limits to the generalisability of the findings. The pattern of solvent abuse among the cases studied here was for the most part experimental or occasional and only a minority had abused solvents frequently. The suggestion that the poor scholastic attainments of solvent abusers can be adequately explained in terms of their domestic circumstances, levels of problem behaviour and antecedent educational attainments cannot necessarily be generalised to cases with more chronic sniffing histories.

ACKNOWLEDGMENTS. — The study is an extension of work carried out at the Department of Clinical Epidemiology and Social Medicine, St George's Hospital Medical School. This earlier phase of the research was carried out with the collaboration of Martin Bland, Senior Lecturer in Medical Statistics, and John Ramsey, Head of the Toxicology Unit at St George's and was supported by a grant from the Department of Health and Social Security to Professor Ross Anderson. The extension of this work to enable the collection and analysis of the pupils' examination results was carried out while the first author was recipient of a Bethlem-Maudsley Scholarship at the Institute of Psychology.

We would like to thank the Directors and Deputy Directors of Education of the participating boroughs, the Research and Statistics Division of the Inner London Education Authority and the Headteachers and Deputy Headteachers of the participating schools for allowing us to conduct the study and for providing data on the pupils' examination results. Our thanks are due to Barbara Maughan for advice on the classification of these results and to Dr Graham Dunn for statistical advice.

Correspondence and requests for reprints should be addressed to Dr O. Chadwick, Department of Psychology, Springfield Hospital, London SW17 7DJ.

REFERENCES

ANDERSON, H. R., MACNAIR, R., AND RAMSEY, J. D. (1985). Deaths from abuse of volatile substances: a national epidemiological study. *Br. med. J.*, 290, 304-307.

BARKER, G. H., AND ADAMS, W. T. (1963). Glue sniffers. *Sociol. soc. Res*, 47, 298-310.

BASS, M. (1970). Sudden sniffing death. *J. Am. med Assoc.*, 212, 2075-2079.

BYFORD, D., AND MORTIMORE, P. (1980). *School Examination Results in the ILEA, 1978*, RS 759/80. London: Inner London Education Authority.

Chadwick, O. (1988). The examination attainments of secondary school pupils who abuse solvents. Unpublished M.Sc. dissertation, University of London.

Chadwick, O., Anderson, H. R., Bland, M., and Ramsey, J. (1989a). Neuropsychological consequences of volatile substance abuse: a population based study of secondary school pupils. *Br. med. J.*, 298, 1679-1684.

Chadwick, O., Anderson, H. R., Bland, M., and Ramsey, J. (1989b). Neuropsychological consequences of volatile substance abuse. *Br. med. J.*, 299, 459 [Letter].

Chadwick, O., Anderson, H. R., Bland, M., and Ramsey, J. (In press). *Solvent Abuse: A Study of Prevalence, Correlates and Neuropsychological Consequences.* New York: Springer Verlag.

Halikas, J. A., Weller, R. A., Morse, C. L., and Hoffman, R. G. (1983). Regular marijuana use and its effect on psychosocial variables: a longitudinal study. *Comp. Psychiat.*, 24, 229-235.

Kandel, D. (1975). Reaching the hard-to-reach: illicit drug use among high school absentees. *Addict. Dis.*, 1, 465-480.

Kandel, D., Davies, M., Karus, D., and Yamaguchi, K. (1986). The consequences in young adulthood of adolescent drug involvement: an overview. *Arch gen. Psychiat.*, 43, 746-754.

Lolin, Y. (1989). Chronic neurological toxicity from volatile substances of abuse. *Hum. Toxicol.*, 8, 293-300.

London Reading Test (1980). Forms A and B: Teacher's Manual. Windsor: NFER-Nelson.

Mahmood, Z. (1983). Cognitive functioning of solvent abusers. *Scot. med. J.*, 28, 276-280.

Maughan, B., Gray, G., and Rutter, M. (1985). Reading retardation and antisocial behaviour: a follow-up into employment. J. Child psychol. Psychiat., 26, 741-758.

Merrill, E. (1985). *Sniffing Solvents.* Birmingham: Pepar.

Nurcombe, B., Bianchi, G., Money, J., and Cawte, R. (1970). A hunger for stimuli: the psychosocial background of petrol inhalation. *Br. J. med. Psychol.*, 43, 367-374.

Robins, L. N., Darvish, H. S., and Murphy, G. E. (1970). The long-term outcome for adolescent drug users: a follow-up study of 76 users and 146 non-users. In Zubin, J., and Freedman, A. (Eds.). *Psychopathology of Adolescence.* New York: Grune and Stratton.

Ron, M. A. (1986). Volatile substance abuse: a review of possible long-term neurological, intellectual and psychiatric sequelae. *Br. J. Psychiat.*, 148, 235-246.

Rutter, M. (1967). A children's behaviour questionnaire for completion by teachers. *J. Child Psychol. Psychiat.*, 8, 1-11.

Rutter, M., Maughan, B., Mortimore, P., and Ouston, J. (1979). *Fifteen Thousand Hours: Secondary Schools and their Effects on Children.* London: Open Books.

Rutter, M. L., Tizard, J., and Whitmore, K. (1970). *Education, Health and Behaviour.* London: Longmans.

Seltzer, C. C., and Oechsli, F. W. (1985). Psychosocial characteristics of adolescent smokers before they started smoking: evidence for self-selection. A prospective study. *J. chronic Dis.*, 38, 17-26.

Sterling, J. W. (1964). A comparative examination of two modes of intoxication — an exploratory study of glue sniffing. *J. crim. Law Criminol. Police Sci.*, 55, 94-97.

Watson, J. (1986). *Solvent Abuse: The Adolescent Epidemic?* London: Croom Helm.

(*Manuscript received* 31*st March*, 1989)

Br. J. educ. Psychol., 60, 192-206, 1990

THE WRITTEN RECALL OF PRINTED STORIES BY SEVERELY DEAF CHILDREN

BY JAMES BANKS
(*Grampian Region Psychological Service, Aberdeen*)
COLIN GRAY
(*King's College, University of Aberdeen*)
AND RONALD FYFE
(*Northern College of Education, Aberdeen*)

SUMMARY. The written recall of printed stories by a sample (N = 16) of severely deaf children (mean age 13:3) was compared with that of a slow-reading hearing sample. The deaf children recalled as much, or more, of the story content. In general, however, their recall contained more distortions of the kind that indicates a break-down of the temporal structure of the story. The writing of one story in Sign word order proved to have a facilitatory effect on cloze recall by the deaf children, but not upon their free recall (as measured by either the amount recalled or the number of distortions), thus clarifying a well-confirmed finding in the literature. The deaf had even more difficulty than the slow-hearing in employing a "top-down", schema-driven strategy at the whole passage level.

INTRODUCTION

THE deaf usually have great difficulty with reading. There is general agreement that "... despite years of effort by teachers to teach them how to read, the vast majority of severely and profoundly deaf children leave school essentially incapable of doing so" (Wood *et al.*, 1986). Surveys of the reading abilities of the deaf have shown consistently that if deaf children with severe prelingual deafness (i.e., with onset of deafness before eighteen months of age) are measured on formal, hearing-standardised tests, they rarely achieve a level of reading proficiency beyond that of a typical nine-year-old child (Trybus and Karchmer, 1977; Conrad, 1979; Wolk and Allen, 1984).

Conrad has argued that our efforts to teach the deaf child to read must inevitably fail, because most deaf children cannot learn the rules of *grapheme-phoneme correspondences* (GPC), which he considers to be essential for word recognition. Since Conrad's gloomy prognostication, however, there have been some developments which have radically altered our way of thinking about the reading of the deaf, so that we have gained at least some understanding of the nature of the problems that they face, if not the key to their solution.

In the study of reading in the hearing, the traditional bottom-up theories have given way to top-down and interactive theories (Stanovich, 1980), according to which the reader actively selects aspects of a passage in accordance with expectations and hypotheses about its overall meaning and structure. In other words, reading can be concept-driven, rather than the meaning emerging in an automatic, stimulus-driven fashion. Accordingly, in the literature on reading in the hearing, there has been a shift in emphasis away from the use of materials consisting of words, parts of words or isolated sentences to meaningful passages, or written discourse, which allow top-down processing full play.

The interactive view has provided helpful interpretations of the difficulties facing the slow-hearing reader (Stanovich, 1980, 1986). For example, in the competent reader, word recognition is so rapid that it is almost certainly bottom-up, rather than top-down (Juel, 1980; Samuels and Eisenberg, 1981). The slow reader, however, may use the immediate context and previous knowledge in top-down fashion as an aid to word recognition. The interactive model thus allows the possibility of compensatory activity, whereby one kind of processing can be deployed to remedy weakness in the other.

There are, no doubt, important parallels between the problems of slow-hearing and deaf readers. Both kinds of reader, for example, tend to "bark at print", that is, they become so entrammelled in their problems with word-recognition and syntax that the general meaning of the passage is lost. Beggs and Breslaw (1983) have shown that typical deaf children read in a word-by-word fashion, rather than showing the selectivity that is characteristic of good readers. As a result of their difficulties, both slow-hearing and deaf readers are subject to the "rich-get-richer"/"poor-get-poorer" mechanism which Stanovich aptly terms the "Matthew effect" (Stanovich, 1986). Recently, there have been attempts to apply interactive models and the Matthew effect to the study of reading in the deaf (Wood *et al.*, 1986; Webster, 1988). In this view, Conrad may have been wrong in supposing that the only route to reading is through the acquisition of the grapheme-phoneme correspondence (GPC) rules. In some circumstances, deaf readers may be able to use contextual information in a top-down fashion to bypass, as it were, their syntactical problems and so arrive at the overall meaning of the passage (Ewoldt, 1981; Cumming *et al.*, 1985).

Yet the deaf appear to face unique problems. Conrad (1979) has produced strong evidence to suggest that deaf readers generally do not have internal speech, which seems to be an essential aspect of competent reading (the very few good deaf readers studied by Conrad had internal speech). Profound deafness may not be an insuperable obstacle to the acquisition of internal speech (Conrad, 1979; Wood *et al.*, 1986), but it is no doubt a major reason for the majority of profoundly deaf children failing to acquire it. And it is a reasonable surmise that in not having an auditory-based internal code, the typical deaf reader may be at a great disadvantage when rehearsing the long stretches of sentence material that must be retained in order to apply the rules of syntax. (Tests of syntactic ability, such as Quigley's TSA (Quigley *et al.*, 1978), show that the deaf have a very poor grasp of syntax.)

But the deaf have an additional problem, as Wood and Webster so cogently demonstrate. The hearing child comes to its formal reading lessons with considerable linguistic knowledge (King and Quigley, 1985). This knowledge has been acquired through constant use of language instrumentally in interaction with its parents and others. It is quite different with the deaf child. Here there is a severe mismatch between the child's linguistic competence and the syntactic sophistication of even the simplest reading matter. As a result, the deaf child finds it difficult to map the reading lesson on to a framework of day-do-day experience.

Almost all studies of reading in the deaf have focused on the reading of single words and isolated sentences; few studies have been made of the reading performance of the deaf on whole passages. This is because the powerful interactive models of reading (Rumelhart, 1977; Lesgold and Perfetti, 1978; Stanovich, 1980), which have aided the understanding of reading processes in the hearing, have only very recently begun to be applied to the study of reading in the deaf. The research reported here was carried out in the belief that the use of interactive models to investigate deaf children's reading of whole passages (stories) might offer insights into the reading problems of the deaf beyond those produced by word and sentence reading tasks. We begin by looking at research into story reading in the deaf and then go on to consider the application of interactive models of reading to their reading problems.

The written recall of printed stories by the deaf

The written recall of stories by the deaf has been little studied. Gaines *et al.*, (1981) compared deaf and hearing children's ability to read a normal English story and two experimentally modified stories. The three stories were equivalent in story-grammatical structure: one was written in ordinary English; another contained words which had been misspelt non-phonetically (e.g. "throgh") as opposed to phonetically ("thru"); and a third contained ambiguous noun and pronoun references (e.g., "He rode on his back all the way to London"). The deaf and hearing children recalled as much of the normal story as the hearing children, but more of the two modified stories. Gaines *et al.* claim that the deaf read "with a broader reconstructive strategy" (p. 467); that is, they make more use of "top-down" processes, such as guessing and reconstructing meaning from partial cues in the text than do hearing children. Since the deaf children's comprehension was not impeded by modifications to textual detail, it was assumed that they engaged in more top-down processing of the modified stories at the whole passage (story schema) level. Additional top-down processing of the immediate context would have resulted in lowered recall, because of the excessive use of cognitive resources for making sense of local detail. The finding that the deaf children employed the same story-grammatical categories as the hearing to organise their recall suggested that they had developed high-level story schemata. The additional finding that the deaf children produced more distortions in their recall was taken as support for the view that a broad, reconstructive, passage-level strategy was being used.

This is an important study, but it has major limitations. Firstly, their sample consisted of six deaf children with a mean reading age of 12.6 years. This level is unusually high for the deaf. Deaf children with more typical (i.e. lower) reading ages would be expected to have greater problems at the level of word recognition than were experienced by the sample studied. Secondly, the sample studied was orally trained. Deaf children who use sign could be expected to have more difficulties than the orally trained at the sentence level, since sign languages such as British Sign Language have been shown to have different sentence organisation (Deuchar, 1983). For example, the sentence "I have the telephone number" would be signed as "Number telephone number have" (p. 70).

In a study by Sarachan-Deily (1985), the abilities of deaf students to recall propositions and make inferences from a story were compared with those of a sample of hearing students. The mean chronological ages of the deaf and hearing samples were 17:2 and 16:1 years, respectively. The two samples had very similar mean IQs, but markedly different reading levels, the mean grade equivalents being 5.4 and 10.5, respectively. The hearing students recalled more story propositions than the deaf students, but the numbers of inferences recalled by the two groups were stated not to differ significantly. Sarachan-Deily suggests that the deaf have impaired propositional recall of stories because of their sentence processing difficulties and argues that they should make use of their background schemata or "world knowledge" about a topic to help them with comprehension. Two points should be noted, however. Sarachan-Deily gives the means and standard deviations of the numbers of inferences recalled by the deaf and hearing students as 4.45 (SD = 1.98) and 6.85 (SD = 2.00), respectively. Since N = 20, t = 3.8138, which has a one-sided P-value of 0.0002. The difference between the means is very substantial; the point biserial correlation is 0.5261 and its square, the correlation ratio, is 0.2768. Thus 27.68 per cent of the variance is accounted for by the difference between the means. There is therefore strong evidence that under the conditions of Sarachan-Deily's experiment, deaf children do *not* recall as many inferences as hearing children. Furthermore, the single story used in her study, "Lost in Alaska", seems to have been easy for *both* groups to read. The implications of this study for the reading process in the deaf, therefore, are unclear.

At the time of writing, no other studies on the written recall of stories by the deaf appear to have been published. It is clear that much more empirical information is required. Further insight into the reading of the deaf may be gained on consideration of models of the reading process and their applicability to the deaf.

Story grammars: do the deaf read "story-grammatically"?

The competent reader is responsive to the emergent features of integrated discourse, as well as to isolated words or sentences. Knowledge of the subject-matter can help the reader to follow the flow of ideas in discourse. In addition, the knowledge of the typical structures of texts can also aid the reading process; in particular, knowledge of typical story structures, stored as "schemata", plays an important role in top-down processing of story (discourse) features. Schemata are organising structures of knowledge, which are acquired through general experience (Anderson, 1978; Rumelhart, 1980). Story schemata are sets of expectations about the internal structures of stories, which guide comprehension and memory during reading and listening. Story schemata have been described by means of "story grammars" (Mandler and Johnson, 1977; Rumelhart, 1977; Stein and Glenn, 1979) which define the elements in the internal structure of an idealised story: *setting*; a *problem*, which the main character must solve; an *outcome*; and an *ending*. Knowledge of such structures, derived from previous experience of story forms, can help the reader or listener to organise and predict incoming story information (Mandler, 1978). Thus, just as knowledge of syntax helps readers to process sentences in a top-down fashion, so knowledge of story grammar helps good hearing readers to process complete stories. Lack of sensitivity to story-grammatical structure has been demonstrated in poor hearing readers (Weaver, 1978, Feagans and Short, 1984). Presumably the focus upon local context leaves no residual cognitive capacity for "story-schematic" reading.

The interactive-compensatory viewpoint may apply *a fortiori* to the deaf reader, whose notorious difficulties with syntax would force even greater dependency upon the use of local context, with the result that they derive even less of the passage's meaning than do slow-hearing readers. The study reported here was designed to investigate this view of the reading of the deaf by comparing them with a slow-hearing group at the same nominal level of reading ability.

METHOD

The three stories used by Gaines *et al.* (1981) were adapted for use in the present study. These stories were apparently simple in content and have been used in previous story grammar research with hearing readers (Mandler, 1978). Two of the stories—"The Unlucky Twins" and "The Penguin and the Polar Bear at the North Pole"—were reproduced exactly as used by Gaines *et al.* "The Unlucky Twins" was written in normal English; it acted as the criterion against which performance on the modified stories could be compared. "The Penguin and the Polar Bear at the North Pole" contained ambiguous pronoun references; additional local contextual processing should be required to establish the referents for the ambiguous pronouns with a consequent loss of overall passage understanding and recall. Since the ambiguous pronoun references should interfere with the global structuring of the story, good subsequent recall should therefore be taken as evidence of compensatory story schema effects. However, processing of individual sentences in this story should not be impaired because the meaning of individual sentences remains intact.

"A Night at the Zoo" was written in sign language (BSL) word order by two profoundly deaf children of deaf parents. These children, aged 12 and 14, were fluent in BSL and had learned it as a first language at home. These "signers" were asked to sign the meaning of each story proposition to each other as naturally as possible. The signed version of each story proposition was translated into an English gloss by a hearing teacher who himself had deaf parents and was fluent in BSL. The English glosses were recorded verbatim and from these a single version of the overall story was produced which was designed to retain as much of the word order of BSL as possible. There is some evidence that passages are easier for deaf people to read when they are rewritten using the same lexical order as sign language (Dawson, 1981). Although most of the evidence relates to word and sentence comprehension, it has been shown that the effect is present when whole passages are rewritten in this way (Odom and Blanton, 1970; Dawson, 1979). Prior to the study, the children's teachers confirmed that the children selected would be equally familiar with the content domain of each story. Thus no systematic differences in content familiarity among the stories were anticipated. Each story had the same story-grammatical structure, comprising a major setting and two episodes, each with a beginning, a response, an attempt, an outcome and an ending.

In an area of research with so few participants available, a within-subjects design, in which the same children served in all the experimental conditions, is necessary. The use of a single story which could be read in more than one form by each child (for example, in English, sign word-order, and with confused pronoun references) would have created carry-over effects from reading to re-reading. To eliminate these effects, equivalent stories were used.

Measuring story reading at the passage and local levels

Recall. Memory for overall passage organisation was measured by written recall. Children's recall protocols were scored for the presence of information in each of the story grammatical categories in order to ascertain their awareness of story structure. In addition, the amount and quality of recall were scored.

Cloze. Cloze passages were used as a measure of top-down processing at the local, or within-sentence, level. Every fifth word was deleted. The hearing literature shows that increasing context beyond five words appears to have no influence upon word predictability (MacGinitie, 1961). Although cloze tests can be contrived to make them sensitive to cross-paragraph information in passages, fifth-word deletion cloze is generally considered to be sensitive to the processing of local contextual constraints within sentences rather than whole-passage organisation (Shanahan *et al.*, 1982). Cloze has been demonstrated to be a valid measure of local contextual processing by the deaf (Odom *et al.*, 1967; Marshall, 1970; Fischler, 1983). However, the deaf tend to score much lower on cloze than the hearing because of their syntactical difficulties (LaSasso, 1978).

Participants

The deaf sample

Sixteen prelingually deaf children, all with severe profound bilateral sensori-neural loss, were randomly selected from a school for the deaf. The sample consisted of 10 girls and six boys.

The method of communication used with the children was *signed English* (Quigley and Kretschmer, 1982), that is, the use of signs to represent the word order and syntax of English. This was intended to help them acquire the structure of English for use in reading and writing. In addition, they were exposed to the natural language of the deaf, *British sign language* (BSL) and were encouraged to use this in everyday communication through interaction with fluent deaf adult signers, who were employed in the school specially for this purpose. Thus all the children in the deaf sample understood BSL. We were assured by teachers that none of the deaf children in our sample had intelligible speech.

The deaf children had been taught to read using materials from reading schemes intended for use with the hearing. Much of the teaching was directed towards helping the children to overcome their difficulties with the traditional "foundation" skills, i.e., GPC rules, word recognition and syntax. This approach has been criticised on the grounds that in general it rests upon an outmoded view of reading itself, and in particular that it ignores the special problems of the deaf (e.g. Howarth *et al.*, 1981).

The hearing sample

Fourteen hearing children (10 boys and four girls) were selected from remedial reading classes in local schools. These children were selected because their reading ages were equivalent to those of the deaf sample on the Young Group Reading Test (Young, 1982). The Young reading test measures word and sentence reading from a given list. An example of a sentence restoration item is "The sky is sometimes — bull, bible, blue, bellow, bald, bluebell". Both samples were therefore matched on word reading skills and some aspects of the handling of sentences. There would be many important aspects, however, in which the two samples would differ in their reading; analysis of errors has shown that the deaf reader adopts very different strategies from the hearing reader, as has been shown by Wood *et al.*

(1981). For example, Quigley's TSA would certainly have showed the deaf sample to be weaker on syntax. The purpose of the Young test was simply to screen the children so that the same story material would be of comparable difficulty for both samples. The samples were also matched on non-verbal reasoning by using Raven's Progressive Matrices (Raven, 1976) from which IQ equivalents were derived. Satisfactory matches for reading age and non-verbal reasoning were achieved, as can be seen in Table 1.

TABLE 1

MEAN CHRONOLOGICAL AGE (C.A.), READING AGE (R.A.) AND I.Q. OF THE DEAF AND HEARING SAMPLES

	N	C.A.	R.A.	I.Q.	Mean Hearing Loss in dB (better ear)
Hearing	14	12.18 (1.83)	8.82 (0.49)	100.3 (11.86)	
Deaf	16	13.25 (1.96)	8.26 (0 69)	109 (10.94)	90.87 (13.67)

Standard deviations in brackets

The slow-hearing children had all received remedial reading classes using a variety of published reading schemes and materials drawn from the various subjects in their curriculum.

Selecting hearing controls: the problem of comparability

In the present study, it was decided to make use of hearing children of normal reasoning ability, but who were slow in their reading development. To date, very little use has been made of slow hearing readers as controls in reading research with the deaf. Where such children have no additional handicaps such as language impairments, subnormal intelligence or dyslexia, they arguably provide hearing controls which are as similar to the deaf children (apart from the factor of deafness) as it is possible to get. According to Stanovich (1980, 1986), slow hearing readers use top-down processing at the word and sentence levels to compensate for their weakness in word recognition. It was therefore felt that this comparison group might be particularly interesting because their reading style might prove to be (among hearing children) most like that of our sample of deaf children. It was expected, however, that analysis of their errors would reveal important differences between the two groups of children.

Procedure

Recall test. Each child was tested individually on every story with the order of presentation of stories counterbalanced across subjects. For the hearing children, the instructions were: "Please read this story as quickly as you can. When you are finished, I will ask you to write down what you remember about the story. Please let me know when you have finished reading." For the deaf children, the same instructions were signed. After completion of reading, written recall was requested immediately. Since pilot trials had shown that the deaf children tended to think that they were being asked for *verbatim* recall, the instructions for written recall were as follows: "Please write down what happened in the story you have just read, as if you were telling it to other children."

Cloze test. The purpose of the cloze test was to measure *meaningful* recall of the story contents at the word and sentence levels. Immediate administration of the test after the written recall test might have prompted verbatim recall of the words used in the stories. The cloze test was therefore administered after a delay of one week. To help the children cue their recall, each child was asked again for written recall of each story (in the original order of presentation). Following the second recall, each child was presented with a cloze version of each story, comprising twenty deletions spaced at five-word intervals (see Appendix 2). The deaf children were told (in sign) that some words were missing in the stories that they had previously read and they were to fill in the spaces so that the stories made sense.

Scoring

Recall test. The written protocols were scored by two independent raters for amount of recall, number of distortions and for information recalled in the different story categories. In terms of amount of recall, each protocol was scored for the presence of story propositions. For each of the 22 propositions in each story, a score of two was awarded for accurate recall, a score of one for a version which, while preserving the general meaning, contained minor distortions of action or character identity, and a score of zero for a serious distortion. In addition, the serious distortions were divided into two categories: (i) major distortions, in which recall was inaccurate and the line of the story was broken. Additions of new and irrelevant material (i.e., confabulations) were also included in this category; (ii) temporal inversions, in which the story elements were themselves correctly recalled but their remembered order of occurrence was wrong. Neither major distortions nor temporal inversions counted towards the recall score.

Inter-rater reliabilities for the minor and serious distortion categories were r = 0.91 and r = 0.93, respectively. On the few occasions on which a temporal inversion was recorded, there was complete agreement between the raters.

Cloze test. Exact replacements of the words used in the original stories were scored one point each. This method of scoring was used because it would appear to be the most objective (Rankin, 1970). Furthermore, allowing synonyms to contribute to the Cloze scores would not have made any difference to the overall pattern of results.

The hypotheses

The general aim of the investigation was to provide comparative data that might offer some insights into how deaf children read whole texts. Three bases for comparison were decided upon. Firstly, the deaf sample was compared with a group of hearing children who were slow readers, who, according to Stanovich (1980, 1986) were also likely to use an interactive-compensatory strategy when tackling a difficult passage. Support for this general hypothesis (1) *that the groups would be similar in the amounts they recall* was sought by comparing performances on the Normal story in terms of a measure of story understanding (written Recall) and of story schema utilisation (story categories used in recall). For both of these measures, *no significant differences between the slow and deaf readers were predicted.* However, the work of Gaines *et al.*, 1981 and the general finding that the deaf seem to have difficulties with successive processing (Cumming and Rodda, 1985) led to the hypothesis that (2) *the deaf sample would produce more temporal distortions within recall than the hearing sample.*

Secondly, the use of two modified stories allowed further differential predictions. It was predicted that the syntactic modifications in the Sign story would benefit the deaf and handicap the hearing. This leads to two testable hypotheses: (3) a "strong effect" hypothesis that *the amount of recall by the deaf of the Sign story would be superior to their recall of the Normal story;* and (4) a "weak effect" hypothesis that *the amount of recall by the deaf of the Sign story would be superior to the amount of recall of this story by the slow-reading group.* Confused pronoun references (as in the second modified passage) make it difficult for any reader to establish a coherent representation of the overall passage meaning. It was hypothesised (5) *that the recall of both the deaf and the hearing groups would be poorer on this story than on the Normal story.* However, evidence from Gaines *et al.* (1981) suggests another interesting hypothesis as an alternative to (5). They argue that, because the deaf have had to learn to adopt a reconstructive, top-down processing strategy to overcome problems in communication, their performance would not be affected (or so badly affected) by these modifications. Logically, this would entail the hypothesis (5a): *the recall by the deaf of the Confused pronouns story would be superior to the recall of this story by the slow hearing group*, provided that the performances of the two groups on the Normal story really are similar, as predicted in the first hypothesis. It is possible, however, that the strategy suggested by Gaines *et al.* may only be used by high reading achievers among the deaf.

The use of the Cloze measure introduces a third basis for comparison. In Cloze, readers are required to reconstruct the passage at sentence and word levels only. The overall structure of the passage is still provided by the remaining text. Deaf children's notorious difficulty with syntax leads to the hypothesis (6): *the cloze scores of the deaf will be inferior to those of the slow group on the Normal story.* However, since the Sign story matches closely the syntax used by the deaf in sign, it can be hypothesised that (7) *the cloze scores of the deaf on the Sign story will be superior to their scores on the Normal story.* The Confused Pronouns story offers a further testable hypothesis: (8) since both groups are hypothesised to operate essentially at the word and sentence levels, *their Cloze performance on the Confused Pronouns story should be similar to their Cloze performance on the Normal story.*

RESULTS

In the analysis of variance of the data, the Group factor (Deaf, Hearing) was analysed as a between-subjects factor, and those of Story Type and Time of Testing were within-subjects factors. With repeated measures factors, it is essential to employ conservative testing procedures, because with non-sphericity of the data, the type 1 error rate may be inflated. The Geisser-Greenhouse procedure, in which statistics of the variance-covariance matrix are used to adjust the degrees of freedom of the final F test, was followed for the test of each within subjects source. The "tail probabilities", or P-values, that we shall report are (for the repeated measures factors) conservative (Geisser-Greenhouse) P-values.

TABLE 2

PERCENTAGE RECALL FOR EACH STORY

	First Recall			Second Recall		
	Normal Story	Sign Story	Confused Story	Normal Story	Sign Story	Confused Story
Hearing	48.2	29.0	32.3	35.5	19.6	19.9
	(15.7)	(19.1)	(14.9)	(16.7)	(11.3)	(12.1)
Deaf	50.5	37.3	28.5	38.2	24.3	20.0
	(24.2)	(26.3)	(21.5)	(26.2)	(20.6)	(22.8)

Standard deviations in brackets

It will be recalled that the deaf and hearing groups were initially matched on non-verbal reasoning and on performance in reading words and sentences. In Table 2 are shown the percentage recall of both groups on each of the three types of story. Both groups performed at much the same level on amount of recall of the Normal story, a measure reflecting to a greater degree understanding of the story as a whole. Analysis of variance showed no significant difference in percentage recall between the deaf and hearing samples ($F = 0.14$; NS). First recall was considerably higher than second recall ($F = 72.34$; $P<0.00005$). In all cases, however, the recall of the children in this study was poor in comparison with those in the samples of Gaines *et al.*, 1981.

The results are consistent with the first hypothesis that the performance of typical deaf readers is similar to that of slow hearing readers. According to the interactive-compensatory viewpoint, however, readers may achieve the same general reading level by different routes. The purpose of gathering comparative data on performance with other reading tasks was to gain further insight into the manner in which the deaf deploy their cognitive resources.

The analysis of variance showed no difference between the deaf and hearing groups on recall of the Sign and Confused Pronoun stories. It can be seen from Table 2, however, that both groups recalled more of the Normal story than they did of either the Sign story or the Confused story. The ANOVA showed the factor of Story Type to have a significant main effect ($F = 42.66$; $P<0.00005$). The basis of the main effect appears to be the difference between recall of the Normal story and recall of the Sign and Confused stories; a Tukey Honestly Significant Difference (HSD) test indicated that recall of the Normal English story was superior to that of either the Sign or Confused Pronoun stories, but that the amounts recalled of the Sign and Confused stories were not significantly different for either group of subjects. This, of course, adds additional confirmation that the two groups were comparable on reading performance. On the other hand, it disconfirms hypotheses (3) (the deaf children should recall more of the Sign story than of the Normal story) and (4) (deaf recall should excel slow hearing recall on the Sign story).

Introducing Sign syntax did not appear to benefit the deaf children as far as amount of recall is concerned. This suggests that easing the deaf child's known problems with syntax does not necessarily improve performance on a global reading task. The hypothesis that for both groups amount of recall would be affected by the introduction of confused pronoun references (5) was supported. There is no evidence that deaf children with more typical levels of reading performance (in the deaf population) adopt the reconstructive strategy that may be adopted by superior deaf readers, such as those studied by Gaines *et al.* In general, then, performance of the two groups was rather similar across the three story types used and so the first basis of comparison did not yield clues as to the nature of characteristically deaf reading, as such. We must turn now to the other bases for comparison.

TABLE 3

MEAN TOTAL DISTORTION OUTPUT FOR EACH STORY

	First Recall			Second Recall		
	Normal Story	Sign Story	Confused Story	Normal Story	Sign Story	Confused Story
Hearing	3.4	2.9	3.2	3.8	3.4	3.6
	(1.5)	(1.8)	(1.9)	(2.5)	(1.2)	(1.9)
Deaf	5.0	4.2	5.9	5.3	4.9	4.5
	(3.4)	(2.6)	(5.1)	(4.7)	(2.1)	(3.0)

Standard deviations in brackets

Distortions

It is clear from Table 3 that the deaf participants averaged more distortions on every story. The analysis of variance showed a significant difference in distortion output between the deaf and hearing groups ($F = 5.49$; $P<0.02$). Further analyses of variance on the different kinds of distortion showed that the deaf sample produced significantly more major distortions ($F=6.99$; $P<0.01$) and significantly more temporal inversions ($F=7.88$; $P<0.008$). Temporal inversions that confuse information from different episodes reflect a more serious breakdown of temporal recall than do inversions of material within episodes. In the hearing sample, only one of the children produced an inter-episodic inversion, whereas in the deaf sample seven children produced inter-episodic inversions. A chi-square test (using Yates' correction) shows significance only beyond the 0.1 level ($\chi^2 = 3.42$; $P<0.10$). On the other hand, there is evidence from Monte Carlo studies that the Yates correction for approximate chi-square tests of association with 2 x 2 tables results in a conservative test (Camilli and Hopkins, 1978). Without Yates' correction, $\chi^2 = 5.12$; $P<0.025$. It would seem, therefore, that deaf children of a more typical level of reading skill (within the deaf population) than those studied by Gaines *et al.* nevertheless do perform more like the superior deaf readers in that study than like the slow readers in the present study, whom they resemble in all other respects discussed so far. Temporal distortions, then, seem to be typical of deaf readers. Hypothesis 2 (the deaf sample should produce more temporal distortions) is therefore supported.

Story-grammatical structuring

The story propositions accurately recalled were assigned to the story-grammatical categories proposed by Mandler and Johnson (1977). Since the number of propositions in each story category was small, the total recall within each category was collapsed across the first and second testings. There were no statistically significant differences between the groups. Once again, however, the performance of both groups was poor. Comparative data provided by Mandler and Johnson show that Grade 2 children, who would have comparable reading ages, reproduced more than 60 per cent of the propositions in all categories other than Reaction (just over 20 per cent) and Ending (approximately 50 per cent). These results suggest that both the deaf and hearing groups make comparatively limited use of story categories in story retrieval. The fact that their performance remains unaffected by the Confused Pronouns passage (where the increased ambiguity reduces the coherence of the narrative) supports this. It is not clear, however, whether this is due to a failure in strategy or to the absence of (or the inability to access) mature schemata.

Cloze

The Cloze data are summarised in Table 4, which gives the deaf and hearing group means and standard deviations for each of the three different stories. Figure 1 shows the plot of cell means against Type of Story. The pattern of the means indicates the presence of a striking interaction between the factors of Group (Hearing, Deaf) and Story Type (Normal, Sign, Confused).

TABLE 4

MEANS AND STANDARD DEVIATIONS FOR CLOZE RECALL OF THE THREE STORIES

	Normal Story	Sign Story	Confused Story
Hearing	12.36 (3.56)	10.71 (1.82)	12.00 (2.08)
Deaf	8.94 (4.02)	11.69 (4.66)	9.06 (4.58)

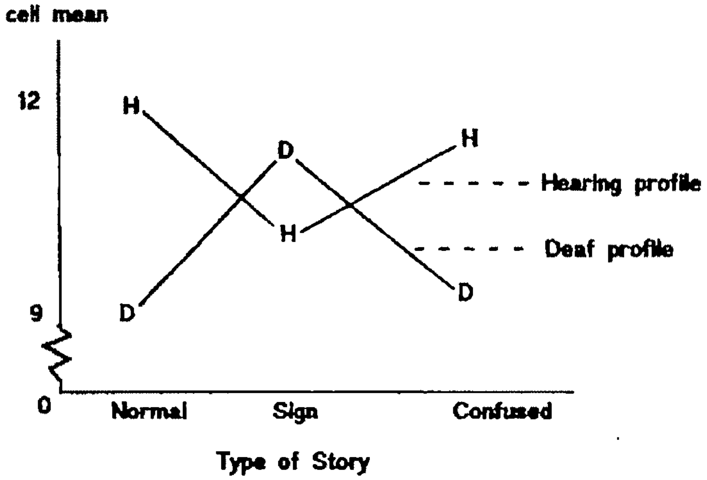
Standard deviations in brackets

With the Normal and Confused Pronoun stories, the hearing group performed better on Cloze; but in the Sign condition, deaf performance was superior. In fact, deaf performance on the Sign cloze was actually superior to their performance on the Normal cloze. The analysis of variance showed this interaction to be significant ($F = 10.29$; $P<0.001$). Before the data were gathered, it was planned to compare the mean deaf sign cloze score with the average of the cloze means for the other two stories. This contrast was significant ($t = 2.66$;

P<0.01). An unplanned, data-snooping Scheffé test (Kirk, 1982) of the same contrast with the data from the hearing group did not show significance.

FIGURE 1

Plots of the Cloze Cell Means against Type of Story for the Deaf and Hearing Samples



In summary, the performance of the deaf children on the delayed cloze passage was markedly different from that of the hearing children. The deaf performed better than the hearing on the Sign story and in addition, their performance on the Sign cloze was superior to their own performance on the Normal cloze. This demonstrates the effect of syntactic problems on the performance of the deaf. More importantly, it confirms that the deaf children focused their efforts at the sentence and word levels. The superior deaf performance on Sign cloze, however, was not reflected in the amount they recalled of the Sign story. Thus, although the performances of the two samples were remarkably similar in many respects, the deaf achieved this despite poor syntactic processing. This suggests that some form of compensatory activity was indeed at work.

DISCUSSION AND CONCLUSIONS

The analysis of errors has shown that the nominal matching (using a simple, hearing-standardised test) of groups of deaf and hearing children on "reading age" can mask fundamental differences in the ways in which they achieve their scores (Wood *et al.*, 1981). The findings of this paper, however, suggest that insights into the reading processes of the deaf can be gained by comparing them with slow hearing readers, from the viewpoint of the interactive-compensatory model, which is proving to be very useful in research with hearing readers and is now beginning to be applied to the reading of the deaf. Broadly speaking, deaf and slow hearing children appear to be similar at the level of total recall of passage content. The deaf, however, have a well-established weakness in syntax (which is reflected here in their performance on the cloze tests) and also in their ability to use story schemata in their reading (as shown by the number and nature of the distortions they produced in their written recall). How, then, did the deaf achieve the same amount of recall as the hearing children? The cloze data for the Sign story establish that under these conditions, deaf children engage in "reconstructive" top-down processing, but at the word and sentence levels they do not read "story-schematically". As a result, they achieve a poor level of recall, similar to that of the slow hearing readers.

An analysis of their errors confirms that the deaf habitually focus on the smaller units of the word and the sentence, thus limiting their residual capacity for gaining access to the global meaning of the passage. This was seen in the kinds of distortion at story-schematic level that were observed in the deaf children's free recall. Why do deaf children adopt this particular strategy, even when, as in the sign-order condition, they receive assistance with their syntactic problems? Recent work by the Nottingham research group (Wood *et al.*, 1986) suggests that many deaf children are exposed to highly directive, controlling teaching, leading to passive learning strategies. Such teaching would be consistent with the failure of deaf children to use story schemata to actively reconstruct global meaning in a story recall task.

Wood *et al.* (1986), Webster (1988), Ewoldt (1981) and others have made several suggestions for the teaching of the deaf. The present study suggests that it may also be useful to give deaf children explicit training in the construction of meaning from whole passages, rather than placing the emphasis solely upon isolated items of vocabulary and single-sentence syntax. Teaching should be directed at helping deaf children to balance the allocation of top-down resources between the whole and local passage contexts.

Future research should concern the nature and role of deaf children's story schemata in their reading generally: is the problem for the deaf one of *strategy* (that is, they have the relevant schemata but fail to use them because of the lower-level processing demands); or is the problem one of deficient knowledge (that is, the deaf fail to acquire story schemata)?

Further research is also needed to investigate the use of compensatory reading strategies in the deaf. In particular, deaf children's failure to use story schemata strategically in their reading suggests that it would be worth attempting to teach story schemata and the strategic use of them to the deaf. It is possible that compensatory processing at the story-schematic level would assist deaf children to reconstruct overall passage meaning and help them to read more "schematically".

Correspondence and requests for reprints should be addressed to Dr James Banks, Grampian Region Psychological Service, 66 Carden Place, Aberdeen, Scotland.

REFERENCES

ANDERSON, R. C. (1978). Schema-directed processes in language comprehension. In LESGOLD, A., PELLEGRENO, J., FOKKEMA, S., GLASSER, R. R. (Eds.), *Cognitive Psychology and Instruction*. New York: Plenum.

ARNOLD, P. (1978). The deaf child's written English — can we measure its quality? *J Br. Assoc Teach. Deaf*, 2, 196-200.

BEGGS, W., AND BRESLAW, P. (1983). Reading retardation or linguistic deficit? III: A further examination of response strategies in a reading test completed by hearing-impaired children. *J. Res Reading*, 6, 19-28.

CAMILLI, G., AND HOPKINS, K. D. (1978). Applicability of chi-square to 2 x 2 contingency tables with small expected frequencies. *Psychol. Bull.*, 85, 163-167.

CONRAD, R. (1979), *The Deaf School Child* London: Harper and Row.

CUMMING, C. E., AND RODDA, M. (1985). The effects of auditory deprivation on successive processing. *Can. J. behav. Sci.*, 17, 232-245.

CUMMING, C. E., GROVE, C., AND RODDA, M. (1985). A note on reading comprehension in hearing-impaired adolescents. *J. Br. Assoc. Teach. Deaf*, 9, 57-60.

DAWSON, E. H. (1979) Cognitive functioning of prelingually deaf children. Unpublished Ph.D. thesis, University of Durham.

DAWSON, E. H. (1981) Psycholinguistic processes in prelingually deaf adolescents. In WOLL, B., KYLE, J., AND DEUCHAR, M. (Eds.), *Perspectives on British Sign Language and Deafness* London: Croom Helm.

DEUCHAR, M. (1983). Is British Sign Language an SVO language? In KYLE, J. AND WOLL, B. (Eds.) *Language in Sign*. London: Croom Helm.

EWOLDT, C. (1981), A psycholinguistic description of selected deaf readers reading in sign language. *Read Res. Q.*, 17, 58-89.

FEAGANS, L., AND SHORT, E. J. (1984) Developmental differences in the comprehension and production of narratives by reading disabled and normally achieving children. *Child Dev.*, 55, 1727-1736.

FISCHLER, I. (1983). Contextual constraint and comprehension of written sentences by deaf college students. *Am. Ann. Deaf.*, 128, 418-424.

GAINES, R., MANDLER, J., AND BRYANT P. (1981). Immediate and delayed story recall by hearing and deaf children. *J. Speech Hearing Res.*, 24, 463-469.

Howarth, S. P., Wood, D. J., Griffiths, A. J., and Howarth, C. I. (1981). A comparative study of the reading lessons of deaf and hearing primary school children. *Br. J. educ. Psychol.*, 51, 156-162.

Juel, C. (1980). Comparison of word identification strategies with varying context, word type and reader skill. *Read. Res. Q.*, 3, 358-376.

King, C. M., and Quigley, S. P. (1985), *Reading and Deafness*. London: Taylor and Francis.

Kirk, R. E. (1982). *Experimental Design: Procedures for the Behavioural Sciences* (2nd edit.). Belmont: Brooks/Cole.

Lasasso, C. (1978). A report on preliminary findings of an investigation of the validity and reliability of the cloze procedure as a measure of readability and comprehension for prelingually, profoundly deaf subjects. In Reynolds, H., and Williams, C. (Eds.), *Proceedings of the Gallaudet Conference on Reading in Relation to Deafness*. Washington D.C.: Gallaudet College.

Lesgold, A. M., and Perfetti, C. A. (1978). Interactive processes in reading comprehension. *Discourse Processes*, 1, 323-336.

MacGinitie, W. (1961). Contextual constraints in English prose paragraphs. *J. Psychol.*, 51, 121-130.

McGill-Franzen, A., and Gormley, K. (1980). The influence of context on deaf readers' understanding of passive sentences. *Am. Ann. Deaf*, 125, 937-942.

Mandler, J. (1978). A code in the node: the use of a story schema in retrieval. *Discourse Processes*, 1, 14-35.

Mandler, J., and Johnson, N. (1977). Remembrance of things parsed: story structure and recall. *Cogn. Psychol.*, 9, 111-151.

Marshall, W. (1970). Contextual constraint on deaf and hearing children. *Am. Ann. Deaf*, 115, 682-689.

Odom, P., and Blanton, R. (1970). Implicit and explicit grammatical factors and reading achievement of the deaf. *J. Read. Behav.*, 2, 47-55.

Odom, P., Blanton, R., and Nunally, J. (1967). Some "cloze" technique studies of language capability in the deaf. *J. Speech Hear. Res.*, 10, 816-827.

Quigley, S. P., Steinkamp, M. W., Power, D. J., and Jones, B. W. (1978). *The Test of Syntactic Abilities*. Beaverton, Oregon: Dormac.

Quigley, S. P., and Kretschmer, R. E. (1982). *The Education of Deaf Children Issues, Theory and Practice*. London: Edward Arnold.

Rankin, E. (1970). The cloze procedure—its validity and utility. In Farr, R. (Ed.), *Measurement and Evaluation of Reading*. New York: Harcourt, Brace and World.

Raven, J. (1976). *The Progressive Matrices*. London: Lewis.

Rumelhart, D. E. (1977). Towards an interactive model of reading. In Dornic, S. (Ed.), *Attention and Performance VI*. Hillsdale N.J.: Erlbaum.

Rumelhart, D. E. (1980), Schemata: the building blocks of cognition. In Spiro, R. J., Bruce, B. C. and Brewer, W. F. (Eds.), *Theoretical Issues in Reading Comprehension*. Hillsdale N.J.: Erlbaum.

Samuels, S., and Eisenberg, P. (1981). A framework for understanding the reading process. In Pizarollo, F. (Ed.), *Neuropsychological and Cognitive Processes in Reading*. New York: Academic Press.

Sarachan-Deily, A. (1985). Written narratives of deaf and hearing students: story recall and inference. *J. Speech Hear. Res.*, 28, 151-159.

Shanahan, T., Kamil, M., and Tobin, A. (1982). Cloze as a measure of intersentential comprehension. *Read. Res. Q.*, 17, 229-254.

Stanovich, K. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Read. Res. Q.*, 16, 32-71.

Stanovich, K. (1986). Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. *Read Res. Q.*, 21, 360-406.

Stein, N., and Glenn, C. (1979). An analysis of story comprehension in elementary school children. In Freedle, R. (Ed.), *Advances in Discourse Processes Vol. 2. New Directions in Discourse Processing*. Norwood: Ablex.

Trybus, R., and Karchmer, M. (1977). School achievement scores of hearing-impaired children: National data on achievement status and growth patterns. *Am. Ann. Deaf*, 122, 62-69.

Weaver, P. A. (1978). Comprehension, recall and dyslexia: a proposal for the application of schema theory. *Bull. Orton Soc.*, 28, 92-113.

Webster, A. (1986). *Deafness, Development and Literacy*. London: Methuen.

Webster, A. (1988). Deafness and learning to read 1: Theoretical and research issues. *J. Br. Assoc. Teach Deaf*, 12, 77-83.

Webster, A. (1988). Deafness and learning to read 2: Teaching Strategies. *J. Br. Assoc. Teach. Deaf*, 12, 93-101.

Wolk, S., and Allen, T. (1984). A five year follow-up study of reading comprehension achievement in a national sample of hearing-impaired students in special education programs. *J. spec. Educ.*, 18, 161-176.

Wood, D. J., Griffiths, A. J. and Webster, A. (1981). Reading retardation or linguistic deficit? 2: Test answering strategies in hearing and hearing-impaired school children. *J. Res. Read.*, 4, 148-156.

Wood, D. J., Wood, H. A., Griffiths, A. J. and Howarth, C. I. (1986). *Teaching and Talking with Deaf Children*. London: John Wiley.

Young, D. (1982). *Group Reading Test*. London: Hodder and Stoughton.

(*Manuscript received 6th March*, 1989)

APPENDIX 1

NORMAL ENGLISH STORY
THE UNLUCKY TWINS

Once there were twins called Tom and Jennifer. They were so unlucky their parents called them the unlucky twins. One day, Jennifer's parents gave her some money to buy the tortoise she wanted. But on her way to the pet shop she lost the money. Jennifer was worried that her parents would be angry with her. So she decided to look everywhere on the pavement where she had walked. She looked in all the cracks on the pavement. She looked in the grass beside the pavement. At last she found the money in the grass. She put the money safely away in her pocket. When Jennifer arrived at the shop, the pet shop man told her that someone else has just bought the last tortoise. He didn't have any more. The same day, Tom fell off a swing. He broke his leg. He wanted to run and play with the other children. So he decided he would not go home. He borrowed a cart from his friends. He asked his friends to pull him around in the cart. While they were playing, Tom fell out of the cart. He broke his arm. Tom's parents said he was even more unlucky than Jennifer. They made him stay in bed until he got well.

CONFUSED PRONOUNS STORY
THE PENGUIN AND THE POLAR BEAR AT THE NORTH POLE

Once there was a polar bear and a penguin. They lived in a house at the North Pole. One day, he got an invitation to work at London Zoo. They said they needed a dancing penguin in their zoo. He thought it would be lots of fun to dance. So he decided to go to London Zoo. He wrote a letter in his best handwriting. He asked them to send the money for a plane ticket to London. Instead of money, they sent a big, grey whale. He rode on his back all the way to London. The penguin became the star of the Zoo. All the children went there to watch him dance. The same day, at the North Pole, a big storm blew down all the animals' houses except theirs. They had to find a place to live. He felt sorry for his friends. The bear wanted to share his house with them. He moved all his furniture outside. He invited them in for the night. The house got more and more crowded. Soon there was no room left for him. The bear had to dress in the bear's warmest clothes. He had to sleep out in the bear's sledge that night.

SIGN WORD ORDER STORY
A NIGHT AT THE ZOO

Every night in zoo, zoo-keeper stay awake watching. Tigers, elephants, and lions outside go play. One night, tigers start tease lions. Tigers say can roar louder than lions. Lions think the tigers stupid. Lions want show can roar loudest in zoo. Lions challenge tigers who win competition. Lions ask monkeys choose who win. Lions made too much noise. Monkeys frightened ran away. Lions never found out who won. All lions hoarse one week. Same night, elephants started fight water fight. Elephants squirt water to other elephants. Zoo-keeper was angry with elephants. Zoo-keeper want elephants stop fighting. Zoo-keeper ran after elephants with big stick. Zoo-keeper shout to elephants get back cages. Elephants squirted keeper. Then elephants ran away. Elephants made elephant yard very muddy. Zoo-keeper cleaned up for two days.

APPENDIX 2

NORMAL ENGLISH STORY
CLOZE TEST

Once there were twins called Tom and Jennifer. They were so unlucky their parents called them the unlucky twins. One day, Jennifer's parents gave her some money to buy the tortoise she wanted. But on the way to the pet shop she lost the money. Jennifer was worried that her parents would be angry with her. So she decided to —— everywhere on the pavement —— she had walked. She —— in all the cracks —— the pavement. She looked —— the grass beside the ——. At last she found —— money in the grass. —— put the money away —— in her pocket. When —— arrived at the shop, —— petshop man told her —— someone else had just —— the last tortoise. He —— have any more. The —— day, Tom fell off —— swing. He broke his ——. He wanted to run —— play with the other ——. So he decided he —— not go home. He borrowed a cart from his friends. He asked his friends to pull him around in the cart. While they were playing, Tom fell out of the cart. He broke his arm. Tom's parents said he was even more unlucky than Jennifer. They made him stay in bed until he got well.

CONFUSED PRONOUNS STORY
CLOZE TEST

Once there was a polar bear and a penguin. They lived in a house at the North Pole. One day, he got an invitation to work at London Zoo. They said they needed a dancing penguin in their zoo. He thought it would be lots of fun to dance. So he decided to go to London Zoo. He wrote a letter in his best handwriting. He —— them to send the —— for a plane ticket —— London. Instead of money, —— sent a big, grey ——. He rode on his —— all the way to ——. The penguin became the —— of the Zoo. All —— children went there to —— him dance. The same ——, at the North Pole, —— big storm blew down —— the animals' houses except ——. They had to find —— place to live. He —— sorry for his friends. —— bear wanted to share —— house with them. He —— all his furniture outside. —— invited them in for the night. The house got more and more crowded. Soon there was no room left for him. The bear had to dress in the bear's warmest clothes. He had to sleep out in the bear's sledge that night.

SIGN WORD ORDER STORY
CLOZE TEST

Every —— in zoo, zoo-keeper stay —— watching. Tigers, elephants, and —— outside go play. One ——, tigers start tease lions. —— say can roar louder —— lions. Lions think the —— stupid. Lions want show —— roar loudest in zoo. —— challenge tigers who win ——. Lions ask monkeys choose —— win. Lions made too —— noise. Monkeys frightened ran ——. Lions never found out —— won. All lions hoarse —— week. Same night, elephants —— fight, waterfight. Elephants squirt —— to other elephants. Zoo-keeper —— angry with elephants. Zoo-keeper —— elephants stop fighting. Zoo-keeper —— after elephants with big stick. Zoo-keeper shout to elephants to get back cages. Elephants squirted keeper. Then elephants ran away. Elephants made elephant yard very muddy. Zoo-keeper cleaned up for two days.

RESEARCH NOTES

Br. J. educ. Psychol., 60, 207-211

THE RELATIONSHIP BETWEEN ADOLESCENT PERCEIVED SELF-COMPETENCE AND ATTENDANCE AT SINGLE-SEX SECONDARY SCHOOL

By Ed Cairns
(*Psychology Department, University of Ulster, Coleraine, N. Ireland*)

Summary. The study examined young people's self-esteem measured using a multidimensional measure of self-esteem, Harter's Perceived Competence Scale, as well as a measure of locus of control, in relation to the type of school they attended — single-sex or co-educational. A sample of students (N = 2295) from both academic and less academic schools was employed. The results indicated that among children (N = 577) attending the more academic grammar schools only, those who attended single-sex schools had higher levels of academic self-esteem and were more inner oriented in terms of locus of control. There were no differential sex effects, however. Additional analyses suggested that this effect was not likely to be due to selection in terms of socio-economic status.

INTRODUCTION

The accepted wisdom that co-education has academic benefits for girls has recently been questioned (e.g. Sarah *et al.*, 1980). Even the earlier suggestion (Dale, 1971) that co-educational schooling might bestow social advantages especially on female students has been strongly challenged (Mahony, 1985).

This anecdotal evidence is now being supported by a growing body of evidence that boys and girls are treated differently in school (Mahony, 1985). In particular there is evidence that girls are made more aware of their academic failings in school. For example Dweck *et al.* (1978) showed that criticism directed at girls by teachers was more likely to attribute poor performance to intellectual failing while criticism directed at boys was of a more general nature and less likely to be related to ability. More recently Frey and Rubble (1987) found that girls made fewer self-congratulatory evaluations of their school performance than boys. Indeed, girls made more critical evaluations compared to boys, thus showing that girls are not only made more aware of their academic difficulties but usually perceive these in a more negative manner.

For some time two variables, self-esteem and locus of control, have been considered to be important in education. In particular self-esteem has been linked to educational attainment in a number of settings (see Byrne, 1984, for a review). Two recent studies (Lee and Bryk, 1986; Foon, 1988) have attempted to examine the impact of attendance at single-sex or co-educational schools upon (among other things) children's self-esteem and locus of control.

Lee and Bryk (1986) carried out a secondary analysis on data gathered in the United States in 1980 which involved 1,807 boys and girls in the second and final years of either single-sex or co-educational secondary schools. Their results, based on regression analyses, led them to conclude that where self-esteem was concerned "the estimated effects generally favoured single-sex schools with larger effects accruing for girls schools" (p. 389). Where boys' schools were concerned they noted that "the picture is more ambiguous".

Foon (1988) reported data from a sample of 1,675 fifteen-year-olds attending co-educational or single-sex secondary schools in Australia. Her method of data analysis involved simply a series of paired-comparison of the four possible groups (boys/girls x single-sex/co-educational school) using t-tests. The results agree with those of Lee and Bryk (1986) in that self-esteem was significantly greater for those pupils attending single-sex schools. However, in contrast to Lee and Bryk (1986), Foon reported that this difference was due to boys at single-sex schools scoring at a higher level than the other three groups. Finally, regardless of type of school attended boys scored at a significantly higher level than did girls, a trend which was also evident in the means reported by Lee and Bryk.

Two criticisms can be levelled at both these studies. The first is that they investigated only a particular type of single-sex school although both tend to make generalisations to all schools. For example the sample employed by Lee and Bryk (1986) consisted of private, Catholic, secondary schools, excluding those with more than 25 per cent vocational students. In other words the schools in their study were those which "emphasise a traditional program". This characteristic of these schools is further accentuated by the fact that over 70 per cent of the students in each of the four types of school (single-sex boys, single-sex girls, co-educational boys, co-educational girls) were identified as "academic track". Less is known about the sample of schools employed by Foon (1988) but it is clear that only non-government schools were involved and that the individual students (95 per cent) rated doing well in school as important or very important, had high-status occupational ambitions, and that their fathers' occupations were reported as being mainly in the high-status categories.

A second criticism concerns the way in which self-esteem was measured. Foon (1988) used as her measure of self-esteem the Piers-Harris Children's Self Concept scale while Lee and Bryk (1986) do not report how self-esteem was measured in their study. However the fact that they report only one score suggests that like Foon they also employed a unidimensional measure of self-esteem. The unidimensional approach to the conceptualisation and measurement of self-esteem is one which has recently been criticised on both theoretical and empirical grounds (e.g., Harter, 1983; Byrne, 1984). Today researchers appear agreed that the self-concept is in fact best viewed as a multidimensional construct (e.g., Marsh and Shavelson, 1985). And in particular it has been suggested (Byrne, 1984) that one of the important constructs which forms part of the self-concept is a specific academic self-concept. Obviously, if this is correct, it is an important construct to measure in research designed to examine the impact of schooling on self-esteem. In adopting a unidimensional approach to the measurement of self-esteem the two studies reviewed above were therefore deprived of information essential to their basic aims.

The multidimensional approach, for example, as advocated by Harter (1983), allows other specific aspects of the self to be measured which may be relevant to the single-sex school debate including ideas about cognitive competence and indeed about social competence. Finally, the unidimensional approach also ignores the idea that in addition to these specific aspects of the self there also exists a general overall self-concept which is relatively independent of the more specific factors (Harter, 1983).

The present study examined data which allow comparisons to be made between boys and girls attending either single-sex or co-educational schools using the Perceived Competence Scale (Harter, 1982), which allows several specific aspects of the self to be measured, including academic self-concept and a more general measure of self-esteem. Further, these data were available for students attending two different types of school in Northern Ireland. These included a sample of children attending Northern Irish grammar schools which teach virtually only academic subjects and where the majority of students stay on beyond the compulsory school leaving age and where the majority go on to university or some other form of tertiary education. The remainder of schools, "secondary schools", have a more vocational orientation entering fewer children for the academically oriented public exams and where the majority of children leave as soon as they have completed their period of compulsory education at age 16.

METHOD

Sample

This was selected from data available on a 1:10 random sample of those pupils in their last year of compulsory education (average age 16 years) in Northern Ireland in 1984 who had taken part in a government commissioned study.

Each of the sample (N = 2,890) was required to complete a lengthy questionnaire while still at school in April, 1984. This questionnaire included, among other things, the Harter Perceived Competence Scale for Children (Harter, 1982). This scale consists of four subscales which measure self-assessed competence in three areas — cognitive, social and athletic (originally labelled physical by Harter) plus a fourth independent scale which measures general feelings of worth or self-esteem. In addition each respondent completed the Nowicki-Strickland Locus of Control Scale (Nowicki and Strickland, 1973) which is a generalised locus of control measure for use with children and young people.

Those who were attending colleges of further education plus those pupils for whom essential data were missing were discarded. This left a total of 2,295 young people (53 per cent girls). Of these the majority (59 per cent) were attending a mixed school. A total of 76 schools were involved, the majority of these (59 per cent) being co-educational (see Table 1).

TABLE 1

Number of Schools (and pupils) in Each Category

	Mixed-sex	Single-sex
Secondary	38 (1088)	21 (630)
Grammar	7 (265)	10 (312)

RESULTS

In order to examine the possible impact of single and mixed sex schools on the perceived competence of boys and girls a two-way analysis of variance, gender x school type (single or mixed) was carried out on each subscale of the Harter test plus the measure of locus of control. Because Northern Ireland has a selective educational system, these analyses were carried out separately for those pupils attending either 17 more academic grammar schools ($N = 577$) or 59 secondary schools ($N = 1718$).

Secondary schools

Boys scored at a statistically significantly higher level than girls on the Cognitive subscale, $F(1,1714) = 18.30$, $P<0.001$, on the Athletic subscale $F(1,1714) = 73.91$, $P<0.001$, and on the measure of General Esteem, $F(1,1714) = 84.98$, $P<0.001$. On the Social subscale girls scored at a higher level than boys, $F(1,1714) = 8.79$, $P<0.001$ (see Table 2). There were no main effects due to the type of school attended and no interactions involving the type of school attended.

TABLE 2

Means (and SDs) on the Harter Test for Secondary and Grammar School Pupils

	Secondary		Grammar	
	Boys	Girls	Boys	Girls
Athletic	2.72 (0.62)	2.47 (0.61)	2.83 (0.67)	2.47 (0.62)
Cognitive	2.63 (0.49)	2.52 (0.53)	2.73 (0.49)	2.61 (0.58)
Gen. Esteem	2.78 (0.48)	2.56 (0.50)	2.78 (0.49)	2.54 (0.50)
Social	2.76 (0.44)	2.92 (0.46)	— —	— —

Grammar schools

Again boys scored at a higher level than girls on the Cognitive subscale, $F(1,573) = 7.98$, $P<0.001$; on the Athletic subscale, $F(1,573) = 40.30$, $P<0.0001$, and on the measure of General Esteem, $F(1,573) = 29.27$, $P<0001$ (see Table 2). There were also two main effects due to the type of school attended with those at single sex school ($X = 2.68$, $SD = 0.52$) scoring at a higher rate than those at mixed sex schools ($X = 2.61$, $SD = 0.51$) on the Cognitive subscale, $F(1,573) = 4.16$, $P<0.05$, and on the Locus of Control measure, $F(1,573)$, single sex = 6.73 (3.39), mixed = 7.50 (3.23).

It is unlikely that these results arise because of a simple cause and effect relationship between the type of school attended and the children's self-esteem. Rather it is possible that these results represent a selection process whereby children with higher levels of academic self-esteem are more likely to attend single-sex grammar schools in Northern Ireland. A limited post hoc test of this idea was possible using the existing data set because demographic information was available for some 87 per cent of the

2,295 young people involved in the analyses reported above (the proportions of boys/girls, grammar/secondary etc. remained close to those reported above). The demographic information available for these respondents was used to produce a measure of social status based on a composite score derived from father's occupation, parents' education and tenure of household. This index of social status was in turn used as a covariate in a replication of the series of two-way analyses of variance reported above. This resulted in no changes in the number of statistically significant differences reported above using the larger data set. This suggests that the selection hypothesis is an unlikely explanation for the results reported here.

DISCUSSION

The results presented above suggest that segregation of children in schools along gender lines has no differential impact on the perceived competence or locus of control of 16-year-old boys or girls who attend such schools. The present results do, however, indicate that being a pupil at a single-sex school may contribute to an increased sense of cognitive competence and a more inner-oriented locus of control. This effect applied only to those young people attending the more academically oriented grammar schools where it applied equally to boys and girls. There was no evidence that attending a co-educational school had advantages in terms of increased social competence for either boys or girls, nor indeed was there any evidence that global self-esteem was influenced by the type of school attended.

In one respect, therefore, the results of the present study replicate both those of Lee and Bryk (1986) and Foon (1988). Both of those studies reported that they had found increased levels of global self-esteem in single-sex schools (although for different groups of students — see above). What the results of the present study suggest is that this effect may be one which is confined to the single domain of cognitive competence rather than having an influence on overall esteem. Further, Lee and Bryk also reported that single-sex schools "delivered" advantages where locus of control was concerned and again this result is replicated in the present results.

A particular contribution of the present study is to confirm that these effects appear to be confined to academic type schools only. There was no evidence that the less academic single-sex schools bestowed any benefits at all on either boys or girls. This is important because the two other studies reviewed above had involved only children attending academically oriented schools.

It is, of course, important to note that the effects reported here are relatively small and indeed confined to one domain of self-perceived competence only and to the area of locus of control. Further, any advantage conferred on girls in terms of cognitive competence by attending a single-sex school was not great enough to overcome the fact that overall girls tended to score at a lower level on this measure compared to boys. Nevertheless the results are important in that they confirm for the third time, and in a third culture, that single sex schools are associated with benefits in terms of self-esteem and locus of control. What the present results do is to refine this finding in terms of the aspects of self involved and the particular type of school involved.

Why academically oriented single-sex schools should function in this way is difficult to understand. Lee and Bryk (1986) make several suggestions. First they consider the possibility that the curriculum in single-sex schools is more academic. It is unlikely that this is true in Northern Ireland where all grammar schools enter students for the same public examinations. A second suggestion is that single-sex schools provide better adult role models because of the greater likelihood of a match between the gender of pupils and those of the staff. The results of the present study make this explanation less likely as this phenomenon is something which almost certainly applies equally to the less academic secondary schools in Northern Ireland. One possibility, therefore, is that this is an important factor but only where students and teachers share educational values and well as gender identity. If this assumption proves to be correct it may help to reinforce the final suggestion made by Lee and Bryk (1986) that "single-sex schools provide not only a positive environment for learning but also a positive environment for teaching". Whatever the explanation, this is an issue which deserves further attention in an era where, in many countries, single-sex schools are becoming an endangered species.

Acknowledgement. — I am grateful to the Policy Planning and Research Unit of the Department of Finance (N.I.) for their permission to use the data reported here and to the Economic and Research Council (U.K.) for financial support.

Correspondence and requests for reprints should be addressed to Ed Cairns, Psychology Department, University of Ulster, Coleraine, N. Ireland.

REFERENCES

Byrne, B. M. (1984). The general/academic self-concept nomological network: a review of construct validation research. *Rev. educ. Res.*, 54, 3, 427-456.

Dale, M. (1971). *Mixed or Single-Sex School/Vol. II. Some Social Aspects*. London: Routledge and Kegan Paul.

Dweck, C. S., Davidson, W., Nelson, S., and Enna, B. (1978). Sex differences in learned helplessness. III The contingencies of evaluative feedback in the classroom. *Dev. Psychol.*, 14, 268-276.

Foon, A. E. (1988). The relationship between school type and adolescent self esteem, attribution style, and affiliation needs: implications for educational outcome. *Br. J. educ. Psychol.*, 58, 44-54.

Frey, K., and Rubble, D. (1987). What children say about classroom performance: sex and grade differences in perceived competence. *Child Dev.*, 58, 1066-1078.

Lee, V. E., and Bryk, A. S. (1986). Effects of single-sex secondary schools on student achievement and attitudes. *J. educ. Psychol.*, 78 (5), 381-395.

Harter, S. (1982). The Perceived Competence Scale for Children. *Child Dev.*, 53, 87-97.

Harter, S. (1983). Developmental perspectives on the self-system. In Mussen, P. H., (Ed.), *Carmichael's Manual of Child Psychology*, Vol. 4, New York: Wiley.

Mahony, P. (1985). *Schools for the Boys· Co-education Reassessed.* London: Hutchison.

Marsh, H. W., and Shavelson, R. J. (1985). Self-concept: its multifaceted, hierarchical structure. *Educ. Psychol.*, 20, 107-125.

Nowicki, S., and Strickland, B. (1973). A locus of control scale for children. *J. consult. clin. Psychol.*, 40, 148-154.

Sarah, E., Scott, M., and Spender, D. (1980). The education of feminists: the case for single sex schools. In Spender, D., and Sarah, E. (Eds.), *Learning to Lose: Sexism in Education*. London: The Women's Press.

(*Manuscript received* 18*th December*, 1989)

Br. J. educ. Psychol., 60, 212-219, 1990

PERCEPTUAL DYSFUNCTION IN POOR READERS: EVIDENCE FOR VISUAL AND AUDITORY SEGMENTATION PROBLEMS IN A SUB-GROUP OF POOR READERS

By Rhona S. Johnston, Marjorie Anderson, David I. Perrett
and Christopher Holligan
(MRC Cognitive Neuroscience Research Group, University of St Andrews)

Summary. The purpose of this study was to consider reading disorders in children within the context of their general perceptual development. A review of the literature suggested that visual and auditory perceptual skills advance from a holistic to an analytical approach, and that reading skills also show this progression. It was proposed that poor readers might show impairments in visual and auditory segmentation tasks, if these skills are important factors in reading development. This hypothesis was tested on a sample (N = 20) of poor readers in the primary age range who were assessed against reading age and chronological age control groups. The hypothesis was partially borne out, a quarter of the sample performing very poorly on both types of tasks. It was argued that segmentation problems in perceptual tasks may underlie reading failure in some children, although the direction of the causation has yet to be determined.

INTRODUCTION

One of the earliest theories advanced to explain developmental reading disorders was the visual perceptual deficit hypothesis (e.g., Orton, 1925). Research in this area, however, has proved equivocal. Some studies have found poor readers to be impaired in copying designs (Koppitz, 1958; Lachmann, 1960; Bender, 1970), or in visual sequential memory (Pumfrey and Naylor, 1978). Other studies have failed to replicate these findings (e.g., Vellutino *et al.*, 1975). Moreover, where evidence favouring a visual perceptual deficit has been found the data have also often been interpretable within a verbal deficit theory of reading disorders (see Vellutino, 1979, for a review).

Visual segmentation skills have, in a somewhat neglected literature, been consistently reported as causing difficulty for poor readers. Goetzinger *et al.* (1960) found that 10- to 12-year-old poor readers were impaired in comparison with chronological age controls on the most difficult subsection of the Gottschaldt Embedded Figures Test, where a shape has to be disembedded from more complex configuration. Lovell *et al.* (1964) subsequently found 10- to 15-year-old poor readers were impaired on the entire task. Additionally, Stuart (1967) found that 12- to 13-year-old poor readers had difficulty with the short form of the Witkin Embedded Figures Test, which is very similar to the Gottschaldt task. However, all of these studies compared poor readers with chronological age controls so it is unclear whether their performance on the embedded figures task was appropriate for their reading age.

It is possible to dismiss poor readers' difficulties with visual segmentation tasks as being due to faulty scanning procedures and attentional problems, as Vellutino (1979) does. However, the importance of auditory segmentation skills in reading (where the child has to extract a sound from a word), and poor readers' problems with such tasks, is well documented. Significant correlations have been discovered between an awareness of the phonemic constituents of words and reading achievement in normal readers (Calfee *et al.*, 1973; Fox and Routh, 1975), and it has been shown that poor readers and illiterate adults suffer impairments on such tasks; (Bradley and Bryant, 1978; Morais *et al.*, 1979; Beech and Harding, 1984; Liberman *et al.*, 1985). However, there may be a reciprocal effect of reading acquisition on the development of both visual and auditory segmentation skills, making firm conclusions about the causal direction uncertain (Gibson *et al.*, 1967; Morais *et al.*, 1979; Read *et al.*, 1986).

Given the evidence from a number of studies that poor readers can suffer from either visual or auditory segmentation problems, it may be productive to consider how children's general perceptual development might relate to success in learning to read. Research on visual perception suggests that children develop from a holistic approach to one where wholes can be segmented into parts (Ames *et al.*, 1953; Ghent, 1956; Vurpillot, 1976). A similar developmental sequence also seems to occur with auditory skills, the young child initially being unable to detect the individual phonemes in a word (Liberman *et al.*, 1974). Gleitman and Rozin (1977) point out that the phoneme is not a natural unit of hearing, so that children have to learn that words can be segmented in this way. These similarities in how auditory and visual segmentation skills develop led Treiman and Baron (1981) to propose that the ability to segment speech may reflect a more general development in the capacity to deal analytically with all kinds of patterns, visual as well as phonological. Such skills are an important element in learning to read, where progress is made through the increasing ability to segment unfamiliar words in order to decode them (Marsh *et al.*, 1981; Frith, 1985). Frith (1985) has suggested that classic developmental dyslexics fail to develop these segmentation skills, and that this is the fundamental problem underlying their disorder.

The purpose of the present study was to examine whether poor readers have impairments in both visual and auditory segmentation ability in comparison with their reading age as well as their chronological age controls, and to determine the relationship, if any, between these skills. Two visual tasks which differ in the extent to which they involve segmentation ability were chosen. One task involved the identification of a shape embedded within a more complex configuration, which was used as an index of visual segmentation ability. This was contrasted with a second task which required the perception of a whole figure whilst ignoring constituent shapes. It was predicted that if the poor readers had visual segmentation problems then they would show impairments in the first task but not in the second. Auditory segmentation ability was also examined using Bradley and Bryant's (1978) odd word out task. It was argued that if deficits in performance on visual as well as auditory segmentation tasks were to be found, then this would imply a general dysfunction in perceptual skills relevant to acquiring competence in reading. Finally, a reading task was carried out in order to assess how the poor readers' word recognition strategies compared with those of their reading and chronological age controls.

METHOD

Sample

Sixty children were studied. Twenty of these were poor readers attending a reading unit several times a week for intensive remedial tuition. Only children with IQs above 90 were selected for attendance at the unit. To be included in the study, each child's reading age on the full version of the British Ability Scales Word Reading Test (Elliott *et al.*, 1977) had to be at least 1.5 years behind their chronological age. The poor readers were matched with both reading age and chronological age controls using the reading test and the short form of the WISC (Maxwell, 1959), the groups being matched for the mean and range of their scores as far as possible. (See Table 1 for means and standard deviations.)

TABLE 1

MEAN CHRONOLOGICAL AGE (CA), MEAN READING AGE (RA), AND MEAN IQ OF THE POOR READERS, READING AGE CONTROLS AND CHRONOLOGICAL AGE CONTROLS (SDs)

	N	CA		RA		IQ	
Poor Readers	20	10.7	(0.5)	7.7	(0.6)	105.3	(15.1)
Reading Age Controls	20	7.5	(0.6)	7.8	(0.4)	107.8	(12.9)
Chronological Age Controls	20	10.6	(0.4)	10.9	(1.1)	107.8	(10.7)

Materials and procedure

All of the testing described below was carried out individually.

(1) *Children's Embedded Figures Test* (Witkin *et al.*, 1971)

This is a measure of visual perceptual disembedding, where the child has to detect and trace out either a triangle or a house shape embedded within coloured pictures of varying complexity. This task explicitly measures segmentation skills.

(2) *Mooney Test* (Mooney, 1957)

This is a test of visual closure and examines the ability to perceive faces. The experimental stimuli consist of black and white pictures in which the removal of half tones accentuates shadows and either obscures or removes detailed features such as the eye. This task differs from the embedded figures task in that it does not require explicit segmentation of the visual configuration, and tests visual synthesis rather than analysis. The children were required to say whether they could see a face and if so to categorise it according to age and sex.

(3) *Auditory Segmentation Task*

This was similar to Bradley and Bryant's (1978) task. There were three conditions, each consisting of eight trials, in which the odd word out had to be detected from a series of spoken words:

(i) Initial phoneme different, e.g., "bud, bun, bus, rug"

(ii) Medial phoneme different, e.g. "dot, cot, pot, bat"

(iii) Final phoneme different, e.g., "fat, sat, pat, bad"

The serial position of the target word was counterbalanced, appearing twice in each serial position. See Appendix 1 for experimental items.

(a) *Memory pre-test*

Prior to carrying out the task, each child was required to recall lists of four words in serial order, so as to assess whether the poor readers were comparable on the memory component of this task with the reading age controls.

(b) *Experiment*

(i) *Practice:* The children carried out practice trials where they had to detect which target word differed from three distractor items in terms of its initial, medial or final phoneme. Any errors in this were explained so that all the children understood the task.

(ii) *Experimental trials:* The trials were presented in blocks, in the order initial, medial, and final phoneme conditions. The words were read out at one second intervals, and the children's responses were noted at the end of each trial.

(4) *Pronunciation of regular and irregular words*

A final task was carried out to assess how the children approached word recognition. According to the dual route model of reading (e.g., Coltheart, 1978), a phonological approach to reading can be indexed by whether there is more accurate pronunciation of regular words (e.g., "best") than irregular words (e.g. "laugh"). Equivalent performance on the two word types, however, would indicate a visual approach.

Regular and irregular words of high and low frequency were selected using Carroll *et al's* Grade 3 norms. The mean word frequencies were as follows: high frequency regular words, mean 324.5 (SD = 216.6); low frequency regular words, mean 27.6 (SD = 24.08); high frequency irregular words, mean 329.0 (SD = 316.2); low frequency irregular words, mean 29.1 (SD = 24.7). These items are listed in Appendix 2.

RESULTS

(1) *Children's Embedded Figures Task*

Each child's overall total per cent correct score was calculated and subjected to an arcsin transformation. A one-way analysis of variance was carried out on these data. The means and standard deviations are to be found in Table 2.

TABLE 2

PERCENT CORRECT PERFORMANCE ON CHILDREN'S EMBEDDED FIGURES TEST (CEFT) AND MOONEY TEST (MT)

	CEFT		MT	
Poor Readers	49.6	(18.2)	68.08	(13.4)
Reading Age Controls	39.6	(15.6)	49.00	(17.0)
Chronological Age Controls	68.8	(12.2)	70.63	(14.4)

Standard deviations in brackets

There was a main effect of groups, $F(2,57) = 18.76$, $P<0.0001$. Newman Keuls tests showed that poor readers were significantly worse than their chronological age controls on this task ($P<0.01$), but that they performed similarly to their reading age controls. The impaired performance by the poor readers in comparison with their chronological age controls seemed to be largely due to a subgroup of children having marked difficulty with the task. In order to quantify this, an analysis was carried out of the number of poor readers who performed more than two standard deviations below the mean of the chronological age controls. A total of eight poor readers were found to be in this category, whereas none of the chronological age controls performed at this level.

(2) *Mooney Test*

The percentage correct score was calculated for each child and subjected to an arcsin transformation. The means and standard deviations are shown in Table 2. A one-way analysis of variance was carried out on these data.

There was a main effect of groups, $F(2,57) = 27.24$, $P<0.0001$, which a Newman Keuls test showed to be due to the reading age controls performing worse than both the poor readers and the chronological age controls ($P<0.01$). No significant difference was found between the latter two groups.

(3) *Auditory Segmentation Task*

(a) *Memory Pre-test.* An analysis was carried out to see if the poor readers differed from the reading age controls in their ability to recall the words correctly according to serial order. A one-way analysis of variance found no significant difference between the two groups, $F < 1$ (see Table 3 for means and standard deviations).

TABLE 3

Percent Correct Performance on Auditory Segmentation Task

	Memory Pre-test	Initial Phonemes	Medial Phonemes	Final Phonemes	Total Correct
Poor Readers	90.95 (19.9)	41.88 (23.7)	57.5 (27.3)	60.0 (17.5)	53.13 (16.8)
Reading Age Controls	86.58 (12.9)	40.63 (21.7)	63.13 (24.4)	58.75 (27.5)	54.17 (17.8)
Chronological Age Controls	96.88 (6.6)	51.88 (25.0)	88.13 (15.4)	84.38 (18.5)	74.80 (12.9)

Standard deviations in brackets

(b) *Experiment.* The percentage of correct responses in each category was calculated for each child and subjected to an arcsin transformation. A two-way analysis of variance with repeated measures was carried out on these data. The between-subjects factor was groups (poor readers, reading age and chronological age controls). The within-subjects factor was conditions (initial, medial and final phoneme). Means and standard deviations are shown in Table 3.

A main effect of groups was found, ($F\,2,57) = 15.25$, $P<0.0001$. Newman Keuls tests showed that the poor readers and reading age controls performed similarly overall, and that both groups were inferior to the chronological age controls ($P<0.01$). There was also a main effect of conditions, $F(2,114) = 28.60$, $P<0.0001$. Newman Keuls tests showed that performance in the initial phoneme condition was significantly worse than in the medial and final phoneme conditions ($P<0.01$), and that performance on the last two conditions did not differ. A significant interaction between groups and conditions was also found, $F(4,114) = 2.88$, $P<0.03$. Newman Keuls tests showed that the poor readers performed similarly on all three conditions, but that the reading and chronological age controls found the initial condition harder than the medial and final ones, there being no significant differences between the latter two conditions. The number of poor readers performing more than two standard deviations below the mean of the chronological age controls was calculated. A total of seven poor readers fell into this category. Overall, five of the 20 poor readers performed more than two standard deviations below the mean of the chronological age controls on both the visual and auditory segmentation tasks.

(4) ***Regularity Task***

The tape recordings were analysed and the percentage of correct responses in each category was calculated for each child and subjected to an arcsin transformation. Due to a technical problem, one poor reader's responses could not be analysed. A three-way analysis of variance for unequal cell numbers was carried out, there being two within-subjects factors, regularity and frequency, and one between subjects factor, groups. Means and standard deviations are shown in Table 4.

TABLE 4

PERCENT CORRECT PERFORMANCE ON REGULARITY TASK

	High Frequency		Low Frequency	
	Irregular	Regular	Irregular	Regular
Poor Readers	75.4	69.4	44.9	54.9
	(19.6)	(19.9)	(22.1)	(26.0)
Reading Age Controls	85.3	85.3	56.5	75.3
	(13.6)	(11.2)	(15.6)	(17.9)
Chronological Age Controls	98.7	99.0	89.6	97.0
	(2.8)	(2.5)	(9.3)	(4.6)

Standard deviations in brackets

A main effect of groups was found, $F(2,56) = 57.18$, $P<0.0001$. Newman Keuls tests showed that the poor readers performed less well than the reading age controls ($P<0.01$), who in turn read less well than the chronological age controls. A main effect of frequency was found $F(1,56) = 137.98$, $P<0.0001$, high frequency words being read better than low frequency words. A main effect of regularity was found, $F(1,56) = 18.61$, $P<0.0001$, but there was also an interaction between groups and regularity, $F(2,56) = 3.77$, $P<0.03$. Newman Keuls tests showed that the reading and chronological age controls read regular words better than irregular ones ($P<0.01$ in both cases), but that the poor readers read both word types equally well. Finally, an interaction was found between regularity and frequency, $F(1,56) = 28.74$, $P<0.0001$. Newman Keuls tests showed that as far as high frequency words were concerned, there was no regularity effect, but that with low frequency words, regular words were read better than irregular ones.

Correlational analyses

A Pearson's product moment correlation test showed that for the poor readers there was a significant relationship between embedded figures and odd word out performance, even when the effects of reading age and chronological age were partialled out, $r(18) = 0.46$, $P<0.05$. The same was the case for the chronological age controls, $r(18) = 0.48$, $P<0.05$, but no such relationship was found for the reading age controls, $r(18) = 0.0009$, NS. Performance on the embedded figures, auditory segmentation and Mooney tasks correlated with reading age in the poor readers, $r(18) = 0.67$, $P<0.01$, $r(18) = 0.49$, $P<0.05$, and $r(18) = 0.57$ $P<0.01$ respectively, the effects of chronological age being partialled out. However, none of these correlations was significant for the controls.

DISCUSSION

The aim of this study was to investigate whether the poor readers were impaired on both the visual and auditory segmentation tasks. They were found to be deficient on both tasks for their chronological age, there being a significant correlation between visual and auditory segmentation ability. No such impairment was found on the Mooney task, which did not tap explicit segmentation skills. Performance on the auditory segmentation task was equivalent to that of the reading age controls (thus replicating Beech and Harding, 1984), as was performance on the visual segmentation task.

Group comparisons assume homogeneity of performance, but an examination of the distribution of the scores showed that not all the poor readers had severe impairments on these tasks. In fact, only five children performed more than two standard deviations below the mean for their chronological age on both the visual and auditory segmentation tasks. However, although the majority of the poor readers performed above these cut-offs, their scores tended towards the lower end of the distribution, there being very few high scorers. For example, on the visual segmentation task only three poor readers performed above the mean of the chronological age controls.

A thorough examination is now needed of how such problems in non-reading tasks relate to slowness in learning to read. If such difficulties cause the reading problems, then the inability to identify familiar elements within a more complex visual configuration may make the recognition and recoding of graphemes into phonemes difficult. Additionally, it may make it hard for the child to recognise frequently occurring sublexical orthographic segments (e.g. "eam" in "beam" and "seam"), and to use these relatively large units as a basis for decoding novel items. Thus these children might be deficient in both alphabetic and orthographic skills, and may show the "logographic" approach to reading that Frith (1985) proposes as being characteristic of the "classic" developmental dyslexic. Some support for this was found in the present study, as the poor readers failed to show an advantage in reading regular versus irregular words, despite the fact that this effect has been found in other group studies of poor readers (e.g., Holligan and Johnston, 1988). It seems likely that because of the depressed visual and auditory segmentation skills in the present group of poor readers, they used an alternative, possibly logographic, approach to word recognition. Such an approach may be more effective for them given their normal performance on the Mooney task, which measured the ability to perceive a whole figure whilst ignoring the constituent shapes.

However, it would be premature to assume that the relationship between these visual and auditory segmentation tasks is a simple causal one. There may in fact be a reciprocal relationship, such that visual and auditory segmentation skills develop in part due to learning to read in an alphabetic script. It would therefore be of interest to determine the extent to which pre-school performance on both of these tasks predicts reading achievement. Although Bradley and Bryant (1983) showed that auditory segmentation skills are predictive of later reading ability, performance on their "odd word out" task accounted for at most 10 per cent of the variance in reading. With so much of reading performance remaining unaccounted for, it is possible that preschool visual segmentation skills also make an important contribution to later success in reading. The tendency to see a dichotomy between visual and auditory skills may prove to be unproductive in considering reading development, where these skills are so closely inter-related.

In conclusion, this study has attempted to demonstrate that it is fruitful to consider reading development within the context of children's general perceptual development, rather than limiting research to the auditory domain. The recent emphasis on phonological skills has led to an impressive literature on the problems experienced by poor readers; however, the present study suggests that we should not overlook the fact that visual segmentation skills may also have an important part to play in learning to read.

Acknowledgements. — The authors would like to thank the staff and pupils in schools in Fife, Dundee and Edinburgh for their co-operation in this study.

Correspondence and requests for reprints should be addressed to Rhona Johnston, MRC Cognitive Neuroscience Research Group, Department of Psychology, University of St Andrews, St Andrews, Scotland.

REFERENCES

Ames, L. B., Learned, J., Metraux, R., and Walker, R. (1953). Development of perception in the young child as observed in responses to the Rorschach Test blots. *J. genet. Psychol.*, 82, 183.

Beech, J. R., and Harding, L. M. (1984). Phonemic processing and the poor reader from a developmental lag viewpoint. *Read. Res. Q*, 19, 357-366.

Bender, L. A. (1970). The use of the Visual Motor Gestalt test in the diagnosis of learning disabilities. *J. spec. Educ.*, 4, 29.

Bradley, L., and Bryant, P. E. (1978). Difficulties in auditory organisation as a possible cause of reading backwardness. *Nature*, 271, 746.

Bradley, L., and Bryant, P. E. (1983). Categorising sounds and learning to read — a causal connection. *Nature*, 301, 419-421.

Calfee, R. C., Lindamood, P., and Lindamood, C. (1973). Acoustic-phonetic skills and reading — kindergarten through twelfth grade. *J. educ. Psychol.*, 64, 293-298.

Carroll, J. B., Davies, P., and Richman, B. (1971). *The Word Frequency Book.* New York: American Heritage.

Coltheart, M. (1978), Lexical access in simple reading tasks. In Underwood, G. (Ed.), *Strategies in Information Processing*. London: Academic Press.

Elliott, C. D., Murray, D. J., and Pearson, L. S. (1977). *The British Ability Scales.* Windsor: NFER-Nelson.

Ellis, A. W. (1985). The cognitive neuropsychology of developmental (and acquired) dyslexia: a critical survey.

Cogn. Neuropsychol., 2, 169-205.

Fox, B., and Routh, D. K. (1975). Analyzing spoken language into words, syllables, and phonemes: a developmental study. *J. psycholing Res.*, 4, 331-342.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In Patterson, K. E., Marshall, J. C., and Coltheart, M. (Eds.), *Surface Dyslexia*, London: Lawrence Erlbaum.

Ghent, L. (1956). Perception of overlapping and embedded figures by children of different ages. *Am. J. Psychol.*, 69, 575.

Gibson, E. J., Farber, J., and Shepela, S. (1967). Test of a learning set procedure for the abstraction of spelling patterns. *Project Literacy Reports*, No. 8, 21-30. (Cited in Gibson, E. J., and Levin, H. (1975), *The Psychology of Reading*. Cambridge, Mass: MIT Press.)

Gleitman, L. R., and Rozin, P. (1977). The structure and acquisition of reading, 1: Relations between orthographies and the structure of language. In Reber, A. S., and Scarborough, D. L. (Eds.), *Towards a Psychology of Reading*. The proceedings of the CUNY conferences. New York: Wiley.

Goetzinger, C. P., Dirks, D. D., and Baer, C. J. (1960). Auditory discrimination and visual perception in good and poor readers. *Ann. Otolaryngol. Rhinol. Laryngol*, 69, 121-136.

Holligan, C., and Johnston, R. S. (1988). The use of phonological information by good and poor readers in memory and reading tasks. *Mem. Cognit*, 16, 522-532.

Koppitz, E. M. (1958). The Bender Gestalt Test and learning disturbances in young children. *J. clin. Psychol*, 14, 292.

Lachman, F. M. (1960). Perceptual-motor development in children retarded in reading ability *J consult. Psychol.*, 24, 427.

Liberman, I. Y., Shankweiler, D., Fischer, F. W., and Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *J. exp. Child Psychol.*, 18, 210-212.

Liberman, I. Y., Rubin, H., Duques, S., and Carlisle, J. (1985). Linguistic abilities and spelling proficiency in kindergartners and adult poor spellers. In Gray, D. B., and Kavanagh, J. F. (Eds.) *Biobehavioural Measures of Dyslexia*. Parkton, MD: York Press.

Lovell, K., Gray, E. A., and Oliver, D. E. (1964). A further study of some cognitive and other disabilities in backward readers of average non-verbal reasoning scores. *Br. J. educ. Psychol.*, 34, 275-279.

Marsh, G., Friedman, M. P., Welch, V., and Desberg, P. (1981). A cognitive-developmental theory of reading acquisition. In Waller, T. G., and MacKinnon, G. E. (Eds.), *Reading Research. Advances in Theory and Practice, 3*. New York: Academic Press.

Maxwell, A. E. (1959). A factor analysis of the Wechsler Intelligence Scale for children. *Br J. educ Psychol.*, 29, 237-241.

Mooney, C. M. (1957). Age in the development of closure ability in children. *Can J Psychol*, 11, 219-226.

Morais, J., Carey, L., Alegria, J., and Bertelson, P. (1979). Does awareness of speech as a sequence of phonemes arise spontaneously? *Cognition*, 7, 323-331.

Orton, S. T. (1925). Word blindness in school children. *Arch. Neurol Psychol.*, 14, 581-615.

Pumfrey, P. D., and Naylor, J. (1978). The alleviation of psycholinguistic deficits and some effects on the reading attainment of poor readers. *J. Res. Read.* 1, 87-107.

Read, C., Zhang, Y., Nie, H., and Ding, B. (1986). The ability to manipulate speech sounds depends on knowing alphabetic spelling. *Cognition*, 24, 31-44.

Stuart, I. (1967). Perceptual style and reading ability: implications for an instructional approach. *Percept. mot. Skills*, 24, 135-138.

Treiman, R., and Baron, J. (1981). Segmental analysis ability: development and relation to reading ability. In Waller, T. G., and MacKinnon, G. B. (Ed.), *Reading Research Advances in Theory and Practice, 3*. New York: Academic Press.

Vellutino, F. R. (1979). *Dyslexia: Theory and Research*. Cambridge, Mass.: MIT.

Vellutino, F. R., Steger, J. A., DeSetto, L., and Philips, F. (1975). Immediate and delayed recognition of visual stimuli in poor and normal readers. *J. exp. Child Psychol.*, 19, 223-232.

Vurpillot, E. (1976). *The Visual World of the Child*. London: Allen and Unwin.

Witkin, H. A., Oltman, P. K., Raskin, E., and Karp, S. S. (1971). *Embedded Figures Test*. Palo Alto, CA.: Consulting Psychologists Press.

(*Manuscript received* 20*th March*, 1989)

APPENDIX 1

EXPERIMENTAL ITEMS FOR AUDITORY SEGMENTATION TASK

(1) *Initial sound different*

bud	bun	bus	*rug*
tip	pig	pip	pit
hid	hit	*lip*	him
peg	*bed*	pen	pet
log	loss	lot	*cod*
pad	man	mat	mad
ham	*tap*	had	hat
roof	room	*food*	root

(2) *Middle sound different*

dot	cot	pot	*bat*
nod	red	fed	bed
fun	gun	*pan*	run
lit	*cat*	bit	fit
name	game	same	*home*
bin	men	hen	ten
feed	reed	*wood*	seed
fish	*mash*	dish	wish

(3) *Final sound different*

fat	sat	pat	*bad*
job	hop	top	pop
pin	win	*sit*	fin
weed	*peel*	need	deed
hard	yard	card	*farm*
moon	loot	hoot	boot
hand	*bank*	sand	land
wig	fig	*rib*	dig

APPENDIX 2

EXPERIMENTAL ITEMS FOR REGULARITY TASK

Irregular words		Regular words	
High frequency	Low frequency	High frequency	Low frequency
both	touch	best	stuck
great	deaf	green	gate
heard	pint	bring	pest
knew	lose	strong	luck
break	aunt	stick	wake
does	bury	door	gang
come	laugh	take	treat
build	doll	dance	base
gone	sew	kept	rub
bowl	soul	turn	mile
love	sword	fine	spade
foot	prove	tune	spear
blood	glove	along	slate
bread	broad	able	dive

Br. J. educ. Psychol., 60, 220-226, 1990

TEACHERS' PERCEPTION OF PRIMARY SCHOOLCHILDREN'S UNDESIRABLE BEHAVIOURS: THE EFFECTS OF TEACHING EXPERIENCE, PUPIL'S AGE, SEX AND ABILITY STREAM

BY MARK G. BORG AND JOSEPH M. FALZON
(Faculty of Education, University of Malta)

SUMMARY. Using a self-administered questionnaire, the attitudes of 844 Maltese primary school teachers towards 16 selected undesirable behaviours were surveyed. Results indicated that pupils' age is a significant moderator of half of the behaviours: these are perceived to be significantly more serious by teachers in the second three years of the primary school than by their colleagues in the first three years Significant boy-girl differences in the seriousness of behaviour were revealed in most of the behaviours as rated by teachers in both the first and in the second years. In addition, length of teaching experience was found to be a significant moderator of the perceived seriousness of some of the behaviours: the less experienced, when compared with the more experienced, judged these behaviours as being more serious. Ability stream did not have any significant effects on the teachers' attitudes towards the undesirable behaviours.

INTRODUCTION

IN a previous paper (Borg and Falzon, 1989) it was argued that generic descriptions of classroom problem behaviour like misbehaviour are of very little help when it comes to drawing up strategies to manage the problem behaviour. Several studies have drawn attention to the dearth of research on the identification of those pupil behaviours which teachers find most troublesome (Merrett and Wheldall, 1984; Wheldall and Merrett, 1988), and on teachers' perceptions of disruptive behaviours in the primary school (Lawrence and Steed, 1986).

Wickman (1928), in a comparative study of the seriousness of children's behaviour problems as perceived by 511 elementary school teachers and 30 mental hygienists in the U.S.A., found that whereas teachers rated misbehaviour in class as the major problem, mental hygienists rated personality and emotional problems as the more serious. Teachers, for instance, rated behaviours like shyness and unsocialness as being less serious in nature than disobedience and destroying school materials; the converse was true for the mental hygienists. Moreover, Wickman found no significant pupil-sex differences in the teachers' judgments of the seriousness of behaviours; that is, no significant differences were observed in the teachers' judgment of the seriousness of behaviours whether manifested by a boy or by a girl. In Israel, Ziv (1970), in a similar study involving 82 elementary school teachers and 45 psychologists, asked subjects to rank each of 30 categories of behaviours in order of seriousness. Results showed that teachers and psychologists agreed that cruelty, dishonesty, aggressiveness and stealing were the most serious behaviours. Borg and Falzon (1989), in a study of the attitudes of 844 Maltese primary school teachers towards undesirable classroom behaviours, found that stealing, cruelty/bullying, and rudeness/impertinence were perceived as the most serious behaviours in pupils. Pupil-sex and teacher-sex proved to be significant moderators of the teachers' attitude towards most of the behaviours. In an investigation of whether teachers' perceptions of problem behaviour are influenced by teacher group ("regular" vs. "special") and classroom context ("disruptive" vs. "non-disruptive"), Safran and Safran (1987) found that the judgments of "regular" teachers tend to be more severe than those of "special" teachers. Results also showed that teachers in "disruptive" classes tended to be less tolerant than their colleagues in "non-disruptive" classes; that is, problem behaviours were perceived to be more serious in "disruptive" classrooms than in "non-disruptive" ones.

The review of relevant literature has shown that factors like length of teaching experience and pupils' age and ability stream have not been considered as potentially significant moderating variables of the teachers' perception of undesirable behaviours in terms of seriousness. The present study, therefore, sought to provide answers to the following questions:

(i) is teachers' perception of the seriousness of undesirable behaviours influenced by the age of the pupils?

(ii) is pupil-sex a significant moderator of the seriousness of undesirable behaviours as perceived by teachers in the first and in the second three years of the primary school?

(iii) is ability stream a significant moderator of teachers' perception of undesirable behaviours?

(iv) is teachers' perception of undesirable behaviours influenced by length of teaching experience?

METHOD

Sample

All teachers in state primary schools in the Maltese Islands who in April 1988 were in charge of a class were invited to participate in the survey by completing a self-administered questionnaire. Of the 1,074 school teachers in post, 844 (78.6 per cent) returned the completed questionnaire; 610 were females and 234 males making up 81.4 per cent and 72 per cent of the population sex-group figures respectively.

Settings

Compulsory school age in the Maltese Islands is from 5 to 16; classes are co-educational in the six years of primary schooling. The primary school years are divided into two three-year blocks (Grades I to III and Grades IV to VI respectively). In the first three-year block children are grouped into mixed ability classes whereas in the second three-year block they are streamed into classes on the basis of performance in national end-of-year examinations in Maltese, English, mathematics, social studies and religion.

Questionnaire

The self-administered questionnaire consisted of two sections the first of which requested respondents to provide information like length of teaching experience and age-group taught (i.e., whether they taught in the first three-year block or the second). Teachers in the second three-year block were further asked to indicate the ability of their class as compared to the average. In the second section teachers were asked to rate each of 16 selected behaviours in response to the question, "If one of your BOY pupils shows this behaviour, how serious a problem would you consider it to be?" Since this study was concerned with the degree of seriousness of behaviour the rating was made on a five-point scale labelled not serious, fairly serious, serious, very serious, and extremely serious. Using the same 16 behaviours and five-point scale, teachers were also asked to rate the undesirable behaviours in response to the question, "If one of your GIRL pupils shows this behaviour, how serious a problem would you consider it to be?" The rating for each behaviour was scored from 0 (not serious) to 4 (extremely serious) to give a measure of the perceived seriousness of that behaviour. The two listings of the 16 behaviours were determined in a random fashion so as to avoid bias as well as response set to the two questions.

Most of the listed behaviours were selected from Ziv's (1970) list which in turn was based on the comprehensive list in Wickman's (1928) pioneering work; others were modified or included with the Maltese primary school context in mind. Moreover, the 16 behaviours were ones with which most Maltese primary school teachers would have been familiar. Since Ziv (1970) had noted that when faced with a large number of items respondents tend to rank with less care it was decided to restrict the list to 16 behaviours.

Procedure

The authors visited all schools to explain to headteachers the purpose of the study and to ask them to distribute the questionnaire in their schools. By means of a covering letter, which also set out the objectives of the survey, teachers were asked to participate by completing anonymously the self-administered questionnaire and return it directly to the authors using an attached stamped addressed envelope.

RESULTS

Table 1 shows the mean rating (together with its standard deviation and rank order) of the seriousness of each of the 16 selected undesirable behaviours as perceived by teachers in the first and second three-year block. Results show that irrespective of age-group taught the two most serious behaviours are: stealing and cruelty/bullying. It is also quite evident that, with the exception of unhappy/depressed, all behaviours are perceived to be more serious by teachers in the second three-year block than by their colleagues who teach in the first three years. This finding is supported by the overall mean ratings, though the difference is not statistically significant. Untidy in personal appearance, lying, disobedience, talkative, cruelty/bullying, stealing, careless/untidy in work and rudeness/impertinence are perceived to be significantly more serious by teachers of the older pupils than by teachers of the younger ones.

TABLE 1

PUPIL-AGE DIFFERENCES IN THE TEACHERS' PERCEPTION OF THE SERIOUSNESS OF BEHAVIOURS: MEANS, STANDARD DEVIATIONS AND RANK ORDERS

	First Three Years			Second Three Years			
Undesirable behaviour	Mean† (N = 393)	SD	Rank Order	Mean† (N = 451)	SD	Rank Order	Difference between means
Untidy in personal appearance	1.19	0.95	13	1.41	1.08	11	-0.22**
Lying	2.09	1.06	5	2.35	1.07	5	-0.26***
Easily discouraged	1.57	0.95	8	1.63	0.95	9	-0.06
Weepy	1.06	1.02	15	1.11	1.05	16	-0.05
Disobedience	2.04	1.09	6	2.37	1.10	6	-0.33***
Talkative	1.04	1.01	16	1.24	1.01	14	-0.20**
Attention-seeking	1.22	1.05	12	1.29	1.04	13	-0.07
Fearful/easily frightened	1.58	1.07	7	1.67	1.07	8	-0.09
Suspiciousness	1.53	1.00	10	1.57	1.01	10	-0.04
Cruelty/bullying	2.60	1.10	2	2.86	1.09	2	-0.26***
Shyness	1.07	0.95	14	1.15	1.00	15	-0.08
Stealing	2.94	1.03	1	3.16	0.95	1	-0.22**
Careless/untidy in work	1.57	0.98	8	1.74	1.03	7	-0.17*
Rudeness/impertinence	2.41	1.07	4	2 67	1 05	3	-0.26***
Unhappy/depressed	2.52	1.12	3	2.45	1.13	4	0.07
Restlessness	1.28	1.33	11	1.37	1.00	12	-0.09
Overall means and SDs	1.73	1.21		1.88	1.22		-0.15

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$

† Not serious = 0; Extremely serious = 4

The mean ratings, standard deviations and rank orders of the attitudes of teachers in the first three-year block towards each of the 16 behaviours when this is manifested by young boys and by young girls are listed in Table 2. The six most serious behaviours in the two pupil-sex groups are: stealing, cruelty/bullying, unhappy/depressed, rudeness/impertinence, lying and disobedience. The most pronounced difference in the mean ratings and rank orders is for untidy in personal appearance (ranked 16th in boys and 11th in girls). With the exception of the mean ratings for stealing and careless/untidy in work, all pupil-sex differences in the teachers' perception of the seriousness of behaviours are statistically significant. Eleven of these behaviours are rated significantly more serious in young boys than in young girls; the converse is true in only three behaviours namely, untidy in personal appearance, lying and disobedience. In addition, the overall mean ratings of the 16 undesirable behaviours show that these behaviours are perceived to be significantly more serious in young boys than in young girls.

Table 3 sets out the mean rating, standard deviation and rank order of each of the undesirable behaviours in boys and in girls as perceived by teachers in the second three-year block. Although slightly different in their respective rank orders, the six most serious behaviours in the two pupil-sex groups are the same as the ones reported for the younger age-group above (namely stealing, cruelty/bullying, rudeness/impertinence, unhappy/depressed, disobedience and lying). Moreover, untidy in personal appearance is the only behaviour for which the difference in its mean ratings and rank orders is particularly pronounced (ranked 16th in boys and 8th in girls). The t-test analysis shows that significant pupil-sex differences in the teachers' attitudes towards the behaviours occur in 11 instances. Untidy in personal appearance, lying and careless/untidy are perceived to be significantly more serious when manifested by older girls than by older boys; the converse is true for the remaining eight behaviours. The overall mean ratings show that teachers of the older age-group perceive these behaviours as being more serious in boys than in girls.

TABLE 2

Pupil-Sex Differences in the Means, Standard Deviations and Rank Orders for Teachers in the First Three-Year Block

	'Young' Boys			'Young' Girls			
Undesirable behaviour	Mean† (N = 393)	SD	Rank Order	Mean† (N = 393)	SD	Rank Order	Difference between means
Untidy in personal appearance	1.01	0.88	16	1.37	0.99	11	-0.36***
Lying	2.03	1.05	5	2.15	1.07	5	-0.12**
Easily discouraged	1.63	0.97	8	1.51	0.92	8	0.12**
Weepy	1.22	1.08	13	0.90	0.94	16	0.32***
Disobedience	1.99	1.11	6	2.10	1.07	6	-0.11*
Talkative	1.09	1.02	15	0.99	0.99	15	0.10*
Attention-seeking	1.31	1.06	12	1.12	1.03	13	0.19***
Fearful/easily frightened	1.70	1.10	7	1.46	1.04	10	0.24***
Suspiciousness	1.56	1.00	9	1.49	1.00	9	0.07*
Cruelty/bullying	2.67	1.10	2	2.53	1.09	2	0.14*
Shyness	1.11	0.96	14	1.02	0.94	14	0.09*
Stealing	2.93	1.04	1	2.94	1.03	1	-0.01
Careless/untidy in work	1.55	1.00	10	1.59	0.96	7	-0.04
Rudeness/impertinence	2.49	1.06	4	2.33	1.06	4	0.16***
Unhappy/depressed	2.56	1.13	3	2.49	1.12	3	0.07*
Restlessness	1.33	1.05	11	1.22	0.91	12	0.11**
Overall means and SDs	1.76	1.20		1.70	1.18		-0.06***

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$
† Not serious = 0; Extremely serious = 4

TABLE 3

Pupil-Sex Differences in the Means, Standard Deviations and Rank Orders for Teachers in the SecondThree-Year Block

	'Old' Boys			'Old' Girls			
Undesirable behaviour	Mean† (N = 451)	SD	Rank Order	Mean† (N = 451)	SD	Rank Order	Difference between means
Untidy in personal appearance	1.10	0.94	16	1.73	1.12	8	-0.63***
Lying	2.31	1.06	6	2.39	1.09	6	-0.08*
Easily discouraged	1.71	0.98	8	1.55	0.91	9	0.16***
Weepy	1.31	1.12	13	0.90	0.93	16	0.41***
Disobedience	2.34	1.14	5	2.40	1.06	5	-0.06
Talkative	1.31	1.03	13	1.17	0.99	14	0.14***
Attention-seeking	1.38	1.05	11	1.20	1.03	13	0.18***
Fearful/easily frightened	1.81	1.11	7	1.53	1.02	11	0.28***
Suspiciousness	1.60	1.03	10	1.54	0.99	10	0.06
Cruelty/bullying	2.94	1.03	2	2.78	1.14	2	0.16***
Shyness	1.23	1.03	15	1.06	0.97	15	0.17***
Stealing	3.19	0.93	1	3.14	0.96	1	0.05
Careless/untidy in work	1.66	1.05	9	1.82	1.01	7	-0.16***
Rudeness/impertinence	2.73	1.05	3	2.61	1.04	3	0.12**
Unhappy/depressed	2.47	1.14	4	2.43	1.13	4	0.04
Restlessness	1.36	1.02	12	1.37	0.97	12	-0.01
Overall means and SDs	1.90	1.23		1.85	1.22		-0.05***

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$
† Not serious = 0; Extremely serious = 4

For the purposes of this study those classes which teachers rated as being above average are here classified as A-stream, the average as B-stream, and below average as C-stream classes. Due to the relatively small number of classes considered by their teachers to be above average, it was decided that the three ability streams should be in fact grouped into two streams: stream A/B and stream C. The means, standard deviations and rank orders of the ratings by teachers in the second three-year block who teach streams A/B and C are shown in Table 4. The four most serious behaviours in the two ability streams are stealing, cruelty/bullying, rudeness/impertinence and unhappy/depressed. Results show that the observed difference in the mean ratings of each of the 16 behaviours in the two ability streams does not reach statistical significance. This is also the case for the difference in the overall mean rating for the two ability subgroups.

TABLE 4

ABILITY STREAM DIFFERENCES IN THE MEANS, STANDARD DEVIATIONS AND RANK ORDERS FOR TEACHERS IN THE SECOND THREE-YEAR BLOCK

Undesirable behaviour	Stream A/B Mean† (N = 266)	SD	Rank Order	Stream C Mean† (N = 185)	SD	Rank Order	Difference between means
Untidy in personal appearance	1.42	1.08	11	1.40	1.10	12	0.02
Lying	2.42	1.10	5	2.25	1.05	6	0.17
Easily discouraged	1.67	0.98	9	1.58	0.91	10	0.09
Weepy	1.10	1.05	16	1.11	1.04	15	-0.01
Disobedience	2.39	1.12	6	2.33	1.09	5	0.06
Talkative	1.24	0.99	14	1.25	1.04	14	-0.01
Attention-seeking	1.29	1.06	13	1.28	1.02	13	0.01
Fearful/easily frightened	1.68	1.10	8	1.67	1.07	7	0.01
Suspiciousness	1.56	1.05	10	1.59	0.97	9	-0.03
Cruelty/bullying	2.86	1.11	2	2.85	1.08	2	0.01
Shyness	1.20	1.05	15	1.07	0.94	16	0.13
Stealing	3.16	0.95	1	3.16	0.95	1	0.00
Careless/untidy in work	1.78	1.04	7	1.67	1.02	7	0.11
Rudeness/impertinence	2.68	1.06	3	2.65	1.04	3	0.03
Unhappy/depressed	2.48	1.15	4	2.41	1.11	4	0.07
Restlessness	1.34	1.01	12	1.42	1.00	11	-0.08
Overall means and SDs	1.89	1.23		1.86	1.21		0.03

* P < 0.05, ** P < 0.01, *** P < 0.001

† Not serious = 0; Extremely serious = 4

For each of the three teaching experience subgroups, Table 5 lists the mean, standard deviation and rank order of the rating of the teachers' perception of the seriousness of each of the 16 behaviours when manifested by pupils. Stealing and cruelty/bullying are perceived as the two most serious behaviours by teachers in each of the three teaching experience subgroups. It is at once evident that there is a trend in how teachers from the three teaching experience subgroups rate the behaviours: for most of the undesirable behaviours the rating decreases (i.e., perceived as being less serious) as length of teaching experience increases. This trend is confirmed by the overall mean ratings; differences between any pair of overall mean ratings, however, are not statistically significant. Comparing first the mean ratings of the behaviours by teachers with less than 11 years experience and by those with 11 to 20 years teaching experience, results show that, with the exception of two behaviours (weepy and shyness), all behaviours are perceived to be more serious by the less experienced subgroup. The t-test analysis shows that the difference in ratings is statistically significant in six behaviours namely disobedience, talkative, stealing, careless/untidy in work, rudeness/impertinence and restlessness. A comparison of the perceived seriousness of each of the behaviours by the least experienced subgroup and by the most experienced one (over 20 years) reveals that of the 16 behaviours only shyness is considered to be more serious by the most experienced teachers (even though the difference does not reach statistical

TABLE 5

MEANS, STANDARD DEVIATIONS AND RANK ORDERS FOR TEACHING EXPERIENCE SUBGROUPS

Undesirable behaviour	Under 11 Years (A)			11 to 20 Years (B)			Over 20 Years (C)			Difference between means		
	Mean† (N = 298)	SD	Rank Order	Mean† (N = 222)	SD	Rank Order	Mean† (N = 324)	SD	Rank Order	A - B	A - C	B - C
Untidy in personal appearance	1.41	1.08	12	1.25	1.01	13	1.25	0.99	11	0.16	0.16	0.00
Lying	2.32	1.05	6	2.17	1.10	5	2.19	1.08	5	0.15	0.13	-0.02
Easily discouraged	1.73	0.97	8	1.57	0.93	8	1.50	0.93	10	0.16	0.23**	0.07
Weepy	1.09	1.05	15	1.15	1.07	14	1.04	1.00	15	-0.06	0.05	0.11
Disobedience	2.38	1.15	5	2.09	1.08	6	2.16	1.02	6	0.29**	0.22*	-0.07
Talkative	1.35	1.09	13	1.03	0.99	16	1.04	0.93	15	0.32***	0.31***	-0.01
Attention-seeking	1.35	1.03	13	1.31	1.09	11	1.13	1.02	13	0.04	0.22**	0.18*
Fearful/easily frightened	1.68	1.07	9	1.66	1.09	7	1.55	1.06	8	0.02	0.13	0.11
Suspiciousness	1.57	1.00	10	1.52	1.00	10	1.55	1.01	8	0.05	0.02	-0.03
Cruelty/bullying	2.83	1.09	2	2.69	1.08	2	2.68	1.12	2	0.14	0.15	-0.01
Shyness	1.09	0.99	15	1.11	0.96	15	1.13	0.99	13	-0.02	-0.04	-0.02
Stealing	3.21	0.96	1	2.99	0.95	1	2.96	1.05	1	0.22*	0.25**	0.03
Careless/untidy in work	1.79	1.02	7	1.57	1.00	8	1.60	1.00	7	0.22*	0.19*	0.03
Rudeness/impertinence	2.70	1.06	3	2.44	1.06	4	2.48	1.05	3	0.26**	0.22**	-0.04
Unhappy/depressed	2.64	1.12	4	2.54	1.11	3	2.30	1.13	4	0.10	0.34***	0.24*
Restlessness	1.48	1.02	11	1.27	0.93	12	1.21	0.99	12	0.21*	0.27***	0.06
Overall means and SDs	1.91	1.23		1.77	1.20		1.74	1.19		0.14	0.17	0.03

* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$

† Not serious = 0; Extremely serious = 4

significance). Nine of the remaining behaviours are perceived to be significantly more serious by the least experienced teachers than by the most experienced ones. The t-test results for the ratings by teachers with 11 to 20 years experience and by the most experienced ones show that significant differences in the perceived seriousness of the behaviours occur in only two instances; that is, attention-seeking and unhappy/depressed are perceived to be significantly more serious by the 11 to 20 years teaching experience subgroup than by those with over 20 years experience.

DISCUSSION

The present results have shown that irrespective of the moderating variable being considered stealing and cruelty/bullying are perceived by Maltese primary school teachers to be the two most serious of the 16 listed undesirable behaviours. With one exception, all the listed behaviours are perceived to be more serious in the older age-group. Interestingly enough, the exception is unhappy/depressed which is seen as being more serious in the younger age-group; perhaps to most primary school teachers, unhappiness and depression appear more inexplicable, and therefore more serious, when manifested in younger children. The present study also suggests that teachers show more awareness of problems in the personality field; this seems to support Ziv's (1970) contention that since Wickman's (1928) study there has been a narrowing of the gap between the attitudes of teachers and psychologists.

The influence of sex stereotypes is evident in the two behaviours which are perceived to be significantly more serious if manifested in girls, namely untidy in personal appearance and lying, while behaviours like weepy, talkative and fearful/easily discouraged are more serious when manifested by boys.

When considering length of teaching experience, results show that teachers become progressively more tolerant of the listed undesirable behaviours; perhaps, as Fields (1986) suggests, experience has taught these teachers that the great majority of undesirable behaviours in the primary school are, at worst, mild.

ACKNOWLEDGMENTS. — The authors wish to express their gratitude to all those anonymous teachers who took part in the survey. They also acknowledge co-operation from the Department of Education and from all state primary school headteachers in the Maltese Islands.

REFERENCES

BORG, M. G., AND FALZON, J. M. (1989). Primary school teachers' perception of pupils' undesirable behaviours. *Educ. Stud.*, 15, 251-260.

FIELDS, B. A. (1986). The nature and incidence of classroom behaviour problems and their remediation through preventive management. *Behav. Change*, 3, 53-57.

LAWRENCE, J., AND STEED, D. (1986). Primary school perception of disruptive behaviour. *Educ. Stud*, 12, 147-157.

MERRETT, F., AND WHELDALL, K. (1984). Classroom behaviour problems which junior school teachers find most troublesome. *Educ. Stud.*, 10, 89-92.

SAFRAN, J. S., AND SAFRAN, S. P. (1987). Teachers' judgments of problem behaviours. *Except Child.*, 54, 240-244.

WHELDALL, K. AND MERRETT, F. (1988). Which classroom behaviours do primary school teachers say they find most troublesome? *Educ. Rev.*, 40, 13-27.

WICKMAN, E. K. (1928). Teachers' list of undesirable forms of behaviour. In *Children's Behaviour and Teachers' Attitudes* (New York, Commonwealth Fund). Reprinted in WILLIAMS, P. (Ed.), (1974), *Behaviour Problems in School*. London: University of London Press.

ZIV, A. (1970). Children's behaviour problems as viewed by teachers, psychologists and children. *Child Dev.*, 41, 871-879. Reprinted in WILLIAMS, P. (Ed.) (1974), *Behaviour Problems in School*. London: University of London Press.

(*Manuscript received 2nd June*, 1989)

BOOK REVIEWS

AARON, P. G. (1988). *Dyslexia and Hyperlexia.* The Netherlands: Reidel, pp. 302, £42.00 hbk. ISBN 1-55608-079-4.

One of the continuing debates in the "dyslexia" literature centres on the issue of whether the term itself services any useful purpose. On the one side are ranked eminently sensible and pragmatic views of Frank Smith, Peter Bryant, Lynette Bradley and many practising educational psychologists who see the dyslexia lobby as clouding the essentially straightforward issue of how best to teach all children with reading difficulties; on the other are dyslexia "experts" who seek ever more novel ways of grouping reading and spelling disabilities into various subtypes which, by definition, require specific forms of intervention and/or treatment. That the millions of words and acres of South American rain forests spent on this debate has hitherto provided no universally satisfactory solution should leave us in little doubt that no such solution will be forthcoming.

Pity, then, the poor reviewer faced with a title such as this. "Dyslexia" poses enough problems, but is at least justifiable at a descriptive level, but what price "hyperlexia"? (Before reading on, have a guess yourself at the putative meaning.) It was with sinking heart, therefore, that I opened this well produced but expensive text. On the whole, however, I found myself quite pleasantly surprised.

The author himself is one of a fairly rare breed of academics who are interested in both research and practical intervention in literacy difficulties from a relatively detached position. His own research over the past decade has been mainly geared towards small scale intensive investigations of groups of college students with unexplained literacy difficulties and he is co-author of a fascinating study of specific reading disability in historical men of eminence. By complementing these studies with investigations of children with reading difficulties referred to a school psychology clinic, Aaron has put himself in a position to write with some authority on this subject.

The book begins with a useful historical overview of the reading disabilities "puzzle", which is placed firmly within the context of the reading process rather than relating it to hypothetical neurological dimensions. At this point the distinction is drawn between "dyslexia", as representing a relative lack of decoding skills in comparison with the ability to comprehend what is written, "hyperlexia", which is essentially the obverse, i.e., advanced decoding skills coupled with poor comprehension, and, thirdly, "non-specific reading difficulties" involving both poor decoding and poor comprehension. Little attention is given to the third group, who, after all, make up the bulk of those with reading difficulties, but a clear and interesting coverage is given to the history and main research findings related to the first two "syndromes". I remain unconvinced of the value of the term "hyperlexia", but it is helpful to have a convenient summary of its use and origins.

Apart from identifying developmental dyslexia as a distinct subtype of reading disability, Aaron sits squarely on the fence with regard to "deficit" vs. "difference" explanations and the issue of heterogeneity vs. homogeneity. In this he provides a refreshing objectivity by presenting evidence from various research studies and discussing their implications. What he does not do, however, is present nearly enough of the evidence for the so-called psycholinguistic approach to reading. Here and elsewhere there is a curiously "dated" feel to the book. The most up-to-date reference to Goodman is 1968, to Frank Smith, 1971. This weakness becomes particularly evident in the chapter on differential diagnosis which plumps for the identification of deficient components of the reading process. Although the author makes sensible criticisms of IQ testing as a means of diagnosing dyslexia, his alternative strategies are not exactly earth-shatteringly new. Most predictably, they lead into a final chapter on treatment and management of reading difficulties which is basically of the "teach dyslexics phonics", "teach hyperlexics by the language experience approach" variety. Where have I heard all this before?

Frankly, anyone looking for the latest word on dyslexia will not be satisfied by this book. Paradoxically, however, it provides one of the better introductions to an area that is a minefield for the uninitiated. It is well worth considering as a beginner's guide despite the price.

R. L. BURDEN

BAKER, C. D., AND FREEBODY, P. (1989). *Children's First School Books: Introductions to the Culture of Literacy.* Oxford: Blackwell, pp. 240, £8.50 pbk. ISBN 0-631-15927-4.

Research into early reading has traditionally given little attention to what actually exists to be read, or to the ways in which texts themselves construct and display a culture's views about literacy, learning

CALCUTTA
CENTRAL

and childhood. Much research has treated reading and writing as neutral technologies whereas they are always learned in particular contexts and embedded in social practices.

This book, and the study it reports, shifts the focus of research towards what the authors call "a beginning reading corpus", a representative body of basal readers (reading schemes). Analysis of 163 books in common use with young children in Australian schools encourages them to speculate that novice readers, in their first encounters with written language in school, are located in particular relationships with book knowledge and, more broadly, with school learning. The first school books are studied from the perspective of linguists who see language as "systemic" and know that cultural and power relationships are always interwoven with linguistic forms and patterns.

There is a patient, well-informed discussion and critique of the notion that school literacy represents a transition from oral to written language. It is this simplistic belief, they claim, which leads to the sanitised language within much of their studied corpus. The rough-hewn nature of children's primary orality is "submerged" in the problem-free, co-operative conversational world of Peter and Jane. They claim that the writers of these kinds of texts are concerned not so much with catching the vitality within everyday narration, but with aiming to socialise children into classroom discourse, text-book language and the power relations therein.

Their close reading into literary theory and cultural studies persuade the authors that texts for the young, however simple-seeming, "construct" their readership. The books here, designed with a specific purpose (school literacy) in mind, also "provide the child readers with a definition of what their identities, interests, attitudes and experiences are conventionally deemed to be". They reveal how such definitions of kinship, gender, childhood itself are propounded. Their skill as researchers (and close readers of texts that do not always grip the imagination for long) is in matching their theoretical perspectives with thorough textual scrutiny. Elaboration in speech, turn taking, the permission sequences between adults and children in the texts, are all shown to play their part in turning linguistic categories into modes of social categorisation. Their analyses of conversational exchanges between Peter and Jane (or, much less often, between Jane and Peter) show a lightness of touch which reminds us that humour is still too rare in educational research.

Their discussion of classroom interactions, and the various modes in which children receive lessons in what they term "the culture of school literacy" is powerful and reverberative. They add to the significant body of ethnographic work by researchers such as Shirley Brice Heath in unpicking the ways in which teachers' "best practices" are often embedded in notions of children's pre-competence (or incompetence). There are some grounded accounts of teacher-pupil discussions of stories in which we see again how children (here, as novice readers) "must know how to do something they are not credited with being able to do in order to participate in the adults' agenda for learning how to do it". Those who come to school sharing the interpretative categories of their teachers are privileged; their accounts, stories, retellings are foregrounded and "count" as reading.

Optimistically, and unusually in such reports of research, there are recommendations which have the tang of innovative classroom practice to them. A more exploratory and all-inclusive approach to books in early years classrooms must be combined with lessons in textuality; in the different ways by which language is crafted and through which "social purposes come to be translated into textual products" (bills, birthday cards, road signs).

The writers are literate about literacy and will encourage their readers to be. The cross-cultural relevance is made plain when we note that the *Ladybird Key Words* and *Breakthrough to Literacy* are within their corpus. I wish that such insights could now be extended to children's literature written for their supposed pleasure, not just the books with pedagogy in mind. That is another research enterprise. The value of this work is in showing us that cognitive theories, or textual analysis, are still inadequate tools for looking at literacy without these writers' insights into contexts, social practices and how *texts* teach.

COLIN MILLS

COLEMAN, A., AND BEAUMONT, J. (1989). *Psychology Survey 7* London: Routledge, pp. 278, £17.95 hbk. ISBN 1-85433-0012-08.

Each volume of the Psychology Survey series inevitably offers diversity. Perhaps Number 7 is the most divergent so far, carrying us to some of the interesting edges of psychology. Annual Reviews Inc.

have yet to include Parapsychology, The Strange Achievements of Idiot Savants and Objective Tests of Freud's theories. In these days when degree courses have become tighter in their orthodoxy of method and content, it is decidedly refreshing to be reminded of the variety of real problems in our domain of enquiry. The description and explanation of the behaviour of idiots savants and of those claiming parapsychological powers are just as legitimate a focus for investigation as is the laboratory behaviour of psychology undergraduates (Occidental strain).

The rare has always been a crucial challenge to theory. So has the Applied. There is no shortage of real life and contemporary issues in this volume: AIDS and children as witnesses are both current concerns. Psychology is certainly contributing to the latter as Graham Davies shows. For the former, I am more suspicious. Keith Phillips sets down what has been done perfectly well. My worry is that the issue exemplifies financial tokenism; governments making known that they are throwing (our) money in the direction of researchers, absolving themselves from the hard decisions and responsibilities—and tempting us to join the game. Hollin and Harvey choose to take us on a Cook's tour of Criminological and Legal Psychology and Sports Psychology; both are sound guides to the developing landscapes.

Some of the topics are so large as to strain the ingenuity of the authors. Annett and Richardson respectively manage to contain their accounts of the states of the art in Skills and Knowledge Representation within their allotted spaces. Bull has to restrict himself to "bodily movement", but retains the title "Non-verbal Communication", a wise decision but it would have been better to narrow the title as well, in the service of accuracy. Likewise, as Lalljee and Widdicombe note, the recently published Handbook of Discourse Analysis runs to four volumes, and the field presents an array too broad. Given that this approach has evoked so much passion from its protagonists and antagonists, I would have preferred to have read about its strengths and weaknesses. That the authors decided to introduce rather than to evaluate is probably correct, and I am merely revealing that I know more about this area than I do about many of the others.

I have no hesitation in confessing that I know least about Chapter I, "Psychological Statistics". It deserves to be chapter I. Macrae does an excellent job in his Guide to Good Statistics. With an introduction that starts with the most elementary issues, he leads the reader quickly to the latest techniques and new ways of thinking. It is excellent to see the 0.05 level put in its proper but limited place. This chapter should encourage more of us to match our problems to their analysis.

Altogether Number 7 is a worthy addition to the series, an interesting variety of introductions, updates and reviews, and not a dull chapter among them.

PETER ROBINSON

DICK, W., AND REISER, R. (1989). *Planning Effective Instruction.* Hemel Hempstead: Prentice-Hall, n.p. ISBN 0-13-679457-2.

WHELDALL, K., AND MERRETT, F. (1989). *Positive Teaching in the Secondary School.* London: Paul Chapman, £6.95 pbk. ISBN 1-85396-075-6.

These two books complement one another, in that while both are addressed to teachers, the one deals with the content and organisation of instruction, the other with its delivery. Both are lucidly written and blessedly brief. Both attempt to provide skills which can be applied realistically within elementary (primary) classrooms without any special equipment or aids and both succeed — up to a point. There is little exhortation and plenty of practical demonstration, but, since Dick and Reiser work at the Florida State University, their book requires more mental adaptation by British readers than that of Wheldall and Merrett, who are at Birmingham University's Centre for Child Study.

Planning Effective Instruction is presented in eleven chapters, each in a standard format. Each chapter includes an Application section that describes activities for a cumulative project. Three themes run throughout the book. The first is that systematic planning underlies effective instruction. The second theme is the diversity of "types of learning outcome" which teachers should recognise and for which, the third theme, a diversity of learning situations should be created. The book draws heavily on the standard texts of Gagné and his associates and provides a useful introduction to them. It is cool and rational, and will be useful as an introductory text in initial teacher training and school-based INSET, especially those concerned with information technology. The trouble, as is frankly recognised, is that the authors "cannot teach you the vast array of complex interactional skills that every teacher must develop in order to work effectively with students in the classroom". *They* think that real teaching begins with the formulation of instructional goals and objectives. Wheldall and Merrett, in their appealing little book, echo this sentiment: "If we are serious about equality of opportunity we must recognise the need for a curriculum appropriate for all our pupils, whatever their cultural or social background." Identifying the appropriate curriculum and retaining the attention of adolescents who prefer to truant has become a key task for the teaching profession. This was brought out in a *Panorama* programme broadcast on 29th January. The truants, newly defined by Right-wing theorists as clients, must be consulted about their interests and aspirations, and teachers must try harder to satisfy these consumers of the product "education".

"Positive Teaching" is the answer. The troublesome behaviour of secondary school pupils, which creates such stress among some teachers, is sternly put into perspective — it is mostly boys who play up, and they indulge in TOOT, HOC and OOS (Talking Out of Turn and Hindering Other Children, Out Of Seat — wandering about) rather than any serious misbehaviours. Classroom management techniques can be based on five principles, which taken together, constitute Positive Teaching:

* teaching is based on the observable
* almost all classroom behaviour is learned
* learning involves change in behaviour
* behaviour changes as a result of its consequences
* behaviours are also influenced by classroom contexts

Most of these prescriptions are commonplace in many secondary schools, and have had to be adopted in the pursuit of GCSE; teachers accustomed to good relationships might find themselves irritated or patronised by being told that they will have to change, and become accustomed to bestowing praise, instead of being so negative. Psychologists running courses might find that they were being placed under some stress themselves if they presented, to an audience familiar with ten years' development on Records of Achievement, such "insights" as: "wall-charts showing tokens of achievement" are good providers of reinforcement. All the same, *Positive Teaching* would be very helpful to some probationer teachers; what is advocated can do little harm, and may well help to encourage more appropriate behaviour through reinforcements which are contingent, immediate, consistent and abundant.

PAUL WIDLAKE

HARGREAVES, D. J. (Ed.) (1989). *Children and the Arts*. Milton Keynes: Open University Press, pp. 164, £9.95 pbk. ISBN 0-33509880-0.

The aim and recurring thesis of this publication is stated at the beginning by the editor, David Hargreaves:

"This book aims to explain some of the processes which underlie children's artistic behaviour and development. Most of its authors attempt to do so from the point of view of developmental psychology, and the crucial question of the relationship between developmental psychology and arts education runs throughout the book. I shall argue that it is essential for these two disciplines to come together. On the one hand, developmental psychologists need to ground their research in the real world of children's arts — in the studio, classroom, rehearsal room or on the stage. On the other hand, arts educators urgently need a developmental basis for their discipline as well as some degree of rigour in their methodology and research."

Hargreaves regrets this has generally not been done and also that the arts are "grossly neglected in relation to other areas of the curriculum" in schools.

Each of the ten contributors (some in collaboration) approaches the area from a different aspect, especially in the central section of the book dealing with children's development within particular artistic domains: drawings; music; writing; sculpture; drama. David Hargreaves and his colleagues at Leicester University provide useful overviews in the first and third parts.

Hargreaves' first chapter on Developmental Psychology and the Arts is especially helpful and takes us straight into some of the crucial issues, including the appropriateness or otherwise of a

Piagetian model and the implications of a "stages of development" theory. He discusses the annexation of the virtues of creativity by those involved in the arts, a claim which he, rightly, regards as an error. Then follows a review of the central literature on the creative development of children and an account of theories of artistic development, many of which are further developed or at least alluded to by contributors later in the book. In particular, the ideas of Gardner receive fairly sustained attention, especially his distrust of the relevance of developmental stage theory, his view that by the age of 7 or so children have accomplished all major symbolic developments, and his conception of multiple intelligences which may develop at different rates and in different ways.

From the evidence cited throughout the book, there are surprisingly similar findings and conceptual agreements between researchers in the various arts and I suspect that it would not be so very difficult to map corresponding patterns of development from these descriptions, even though they are made by different observers of children's artistic production from apparently different theoretical perspectives. Unfortunately, though inevitably, the terminology that these writers employ tends to be disparate. However, reading, so to speak, between the lines, it seems possible to identify broad conceptual agreement between, for example, Parsons, Wolf, Bolton and Swanwick; though doubtless some of these might wish to dissociate themselves from the views of others.

I must confess to finding myself still slightly puzzled by the approach taken by my colleagues Davidson and Scripp on education and development in music "from a cognitive perspective". Their particular view emanating from *Project Zero* seems to focus fairly sharply on the idea of "representation", rather along the lines developed by Bruner, and representation in this case is clearly envisaged to be through various forms of notation. This seems a particular difficulty in music where notation, if used at all, is a *secondary* activity; unlike procedures in the visual arts where, for example, the "notation" of a drawing may be in black and white, in blocks of colour and so on. The representational process in music is surely articulated in sound and thus should be studied through sonorous objects, the music children make. Perhaps it is the cultural context of a largely performance-based American school and college curriculum which to some extent has determined this particular analysis of musical development through traditionally or graphically notated music. Things are different elsewhere, for instance in Britain and Scandinavia, where the strong movement towards children *composing* points towards a very different research paradigm and opens up a range of data not so easily accessed in the USA. (In a recent study by this writer, around 18 per cent of available lesson time in secondary schools was committed to individual, small group or whole class composing.)

Children and the Arts is certainly centred on developmental psychology and the artistic products of children. What seems to be missing is any consideration of a theory of arts criticism. It is salutary to challenge developmental theory with the results of observations in "natural" settings, and these observations themselves will modify and create better theories. It would still seem necessary though to develop a third strand; to attempt a conceptual map of artistic experience linked to artistic criticism, running in a continuum from the work of the youngest children to the work of practising artists. The much maligned APU document, *Aesthetic Development*, did at least make a stab at this; though its force was lost in the way in which the report was finally and hurriedly assembled.

Without any question, *Children and the Arts* will be found helpful, drawing our attention to a range of work issuing from a variety of perspectives and clearly setting out for us some of the important themes.

KEITH SWANWICK

HOWE, M. J. A. (1989). *Fragments of Genius*. London: Routledge, pp. 178, £25.00 hbk. ISBN 0-415-00781-X.

This is a fascinating short introduction to the curious phenomena of the "idiot savant". As the author points out, the term "idiot savant" is somewhat unfortunate, but we are stuck with it. In a televised QED programme it was translated as "foolish wise one", which is perhaps even more misleading. "Idiot" in French means "imbecile" or "natural" (in a now almost obsolete sense), while a "savant" is a scientist or scholar. So an idiot savant perhaps translates as an imbecile scholar. Just to add to confusion, the condition has recently become part of popular mythology, no doubt partly due to the film *Rain Man* starring Dustin Hoffman, with the result that mass media now refer to "savants" for short, without explanation.

Howe's title, on the other hand, is a very good one, since the things that these otherwise retarded individuals can do, do precisely resemble bits of what, in a more integrated and complete form, might well attract the label "genius". As is well known, their accomplishments include playing musical instruments and (to a lesser extent) composing; drawing and painting; many remarkable feats of memory; equally remarkable feats of calculation, both arithmetical and calendrical; and a number of other activities. It does not appear to be the case that there is any recorded feat that could not be achieved by a "normal" individual, but there are many that are not normally achieved. It should be remembered that both terms of the label "idiot savant" refer to groups of dimensions, rather than to all-or-nothing conditions. This is only what is true of behaviour in general. Thus "idiocy", even if operationalised as "low measured intelligence", is a characteristic of separate groups, two main ones being autism and other causes such as brain damage; and intelligence is conceived as continuous. Similarly the feats of the individual may be more or less remarkable and as already mentioned occur in various modes. It is naturally the more unusual individuals who attract notice.

Howe begins with general introductory points similar to these, and then describes two cases in some detail. He then considers possible explanations for feats of memory, followed by the question of the independence of human skills. Then come calendar calculating and visual arts, and finally there is a particularly interesting chapter (to me) on extraordinary mental feats in general. The thrust of Howe's argument is, first, that these somewhat bizarre phenomena can be explained, and moreover explained in the same sort of way as is applied to more common behaviour. Thus, in particular, feats that require extensive practice do not appear magically without it (as some of the popular literature has implied), whatever the individual's level of intelligence. It is however necessary to account for the fact that some individuals will devote extraordinary effort to what for most would be an utterly boring activity, such as memorising dates or the details of bus routes. And the same explanation will not necessarily fit different cases.

Second, Howe sees the phenomena as supporting what is now the fashionable modular view of human abilities. Whether this is more than a fashion time will perhaps tell. Certainly there seems to be a return to a more-or-less pre-Galtonian "faculty" view; and the emphasis on practice can be seen in the same way — it will be recalled that Darwin held, prior to reading *Hereditary Genius*, that apart from extremes of genius and idiocy, individual differences in achievement were to be explained by effort rather than ability. Howe does not mention K. Anders Ericsson, the leading current exponent of the "practice makes perfect" view, and he cites only one of several important studies by N. O'Connor and B. Hermelin. But he stresses that he has not attempted an exhaustive survey, and given a limit of less than 200 pages it would be quite unreasonable to expect one. Two minor complaints are that £25 is a lot for such a quantity (no paperback price is given), even now; and that the proof reader has allowed several errors to slip past.

But these are small negative points in a book which should be of great interest both to many psychologists and to the general reader. It is written at a level which does not presuppose technical knowledge, but is nevertheless informed by psychological theory and research. To my knowledge it is only the second book to appear on the subject, the first being D. A. Treffert's *Extraordiary People* (Bantam Press, 1989), and despite its brevity this is perhaps the better introduction. It deserves to be widely read.

John Radford

Langford, P. (1989). *Children's Thinking and Learning in the Elementary School.* Lancaster, PA; Technomic Publishing Co., pp. 171, n.p. ISBN 0-87762-604-9.

What this book clearly is, is very readable. It is also very helpful. What it may be somewhat less, is right up-to-date. For instance, on language, I'd like to have seen a little of Wells; on reading, of Raban; on maths of Desforges; on communication and thinking, something of Wood or Robinson. But perhaps I am being ethnocentric and perhaps I am forgetting that the main audience (judging from both the spelling and selection) is American and Australian.

What is here is very accessible, logically developed and clearly written. Moreover, the author has a direct, engaging, no-nonsense style which should make this a thoroughly useful contribution to initial teacher-training. The slender introduction manages to look at the work of Piaget, Bruner, Gagné and Ausubel in a fresh, useful and careful way. Langford is not dismissive, despite the sweep of his canvas.

(Bruner wins by a comfortable lead in the relevance of his ideas to schooling.) I don't always find myself in agreement with the author: "the most influential version (of process writing) in elementary school has been that put forward by the American educator Donald Graves". But there is much to be gained from this chapter on writing; it is both instructive and enjoyable.

It was at this point that I recognised why the book unfolds so smoothly. Unlike most education books, this is written by a person who can tell a good story and who does not use six words where one will do. At one point Langford says, "Fruitful learning involves a balance between the efforts, discoveries and interests of the learner and the guidance, suggestions and enthusiasm of the teacher" (p. 50). Not earth-shattering, maybe, but certainly a comment worth extracting and sending to the National Curriculum Council, currently preoccupied with terms like "delivering knowledge"!

The writer's enthusiasms come over well, particularly here and in the subsequent chapter on painting and drawing. In that fourth chapter, as in much of the book, Piagetian theory is the leit-motif of much of the analytical schema presented. But Langford, whilst employing Piaget's ideas, often uses them critically and sets other ideas against them. He points out the weaknesses in Piaget's "the child draws what he knows", citing work by Booth, Freeman and Van Sommers which clearly demonstrates other and more convincing modes of analysis. The vigorous and well-chosen illustrations enliven this chapter, too, as Langford moves on to discuss teaching methods for art programmes. His blend of analysis, explanation and guidance strikes me as particularly well done in this section.

Something of the flavour of this book reminds me of much earlier works by de Lissa, Isaacs or Hadfield. Analysis and reflection is here in good measure, but the author is not afraid of being usefully prescriptive.

Overall, then, there is much which is helpful. Who would quarrel with the author's plea for better and more systematic exposure to science for elementary children? He even provides the beginnings of a syllabus! I liked his conclusions here, too. "Now that we have been through the experience of purely child-centred education and realise that children do need adult guidance and even adult inspiration, there is still no reason to go back to a purely adult-centred curriculum" (p. 129). Whilst some of us might doubt whether very many *did* go through that experience, Langford has surely achieved the right balance. That is, for me, the attraction of the whole book; a balanced diet, attractively laid out and enjoyable to consume. Student teachers could be fed some of this to their advantage.

Philip Gammage

Jones, N. (Ed.) (1989). *School Management and Pupil Behaviour*. Basingstoke: Falmer Press, pp. 277, £22.00 hbk., £9.95 pbk. ISBN 1-85000-591-5; 1-85000-592-3.

This book edited by Neville Jones is one of the "Education and Alienation" series from the Falmer Press of which he is the general series editor. The whole range, of about ten books, are of a consistent high quality and this particular one being reviewed must be one of the best.

Neville Jones needs little introduction as he has written widely on SEN and educational psychology issues and this experience shows itself in the careful selection of contributors and their respective chapters. Many of the writers are well known but their material is no mere re-hash of previous publications. A common problem with books such as this is that experienced professionals tend to perceive value and enlightenment from a few chapters only but that is certainly not the case with this book.

It is suggested within that this volume can act as an accompaniment to the Elton Report and this it does admirably by supporting and extending the debate. Used in this way it is an ideal recommended book for the many "priority" and INSET courses currently focusing upon discipline in schools.

The book is organised into three main parts; part one, "Pupil behaviour and school discipline" includes work on effective schools and pupil behaviour by David Reynolds and three other chapters on approaches to disruption including one of the high quality expected of Delwyn Tattum. The second section details six submissions by leading bodies to the Elton Committee. The third section usefully discusses pupil behaviour in the context of curriculum approaches. Of the five chapters, the first one by Richard Pring questions the likely effect of the National Curriculum not least in the light of developing approaches to the curriculum that were more pupil-centred, relevant to low-attainers and the potentially disaffected and accommodating. His discussion succinctly encapsulates this concept.

Neville Jones argues for a new role for educational psychologists in the final chapter of this section and this is worth quoting: "There is a need for EP services to acquire a new status and a new role . . . avenues for this work are through the *ordinary* channels of effective school management, inspiring teaching, developing appropriate curriculum, and the development of appropriate value systems." No mean task.

The epilogue, also written by Neville Jones, is where the most detailed and crisp discussion of the Elton Report takes place.

In the preface Roy Atkinson reports a criticism of the Elton Report that "it is full of bland exhortations and platitudes. We concluded that we had to, in some instances, restate the obvious because we found that some schools were not following what seemed to us obvious good practice in simple matters". But despite the utility of the Elton Report, and the recent spawning of several books on pupil behaviour, the problem remains, as Reynolds reminds us, in how to put well-established educational principles into effective practice.

This book, which I would go as far to recommend as essential reading as an accompaniment to the Elton Report, could be effective in starting to bridge that gap.

M. I. Smith

Lomax, P. (1989). *The Management of Change*. Clevedon, Avon: Multilingual Matters Ltd., pp. 183, £8.95 pbk. ISBN 1-8539-060-6.

This edited collection of papers consists of accounts by individuals who have engaged in changing an aspect of educational practice. All the contributors describe themselves as teachers in different institutional contexts, and all point to the value of action research as the method which has enabled them to become authors or facilitators of change.

What is enlivening and refreshing about the book is the style in which the various case studies are written. The language and form of the accounts, whilst unusual in adopting the first person, are immediately accessible. Furthermore the reader is given an understanding not just of the research process in relation to a range of substantive issues, but insight into the values which have helped inform the teachers' reflections, actions and theorising. Dialogue and poetry are used to capture the richness, feelings and practices of those involved, which is a stark contrast to the way research findings are conventionally articulated and reported. As such the presentation accords with the title of this new series of occasional papers from BERA: "Dialogues".

The book is divided into three parts. The first provides an overview of what is involved in action research, its distinctive cycles of reflection, action, and evaluation, and what its exponents see as its particular relevance to teachers. The contributors are all zealous in their praise for action research. Barrie Jones in an original and stimulating account of his own conversion to action research, does indeed append to himself albeit tongue in cheek "the born again" epithet. Action research is not, however, depicted as unproblematic. It is Mary Gurney's article which highlights some of the difficult political questions facing action researchers whose own agenda for change bring them into conflict with institutional hierarchies.

Part two presents three case studies of the use to which action research has been put in schools and its particular contribution to the professional development of teachers and to school effectiveness. These case studies deal with attempts to change attitudes towards gender equality in a special school, to develop co-operative teaching in an infant school, and to improve classroom interaction in a secondary school. The studies are very much in line with approaches currently being advocated by the Department of Education and Science regarding planning for school improvement. In that they show how such an approach is clearly possible they should hearten others who might wish so to engage.

The concluding section points to those who, as in-service providers, have helped teachers initiate action research and raises questions as to the need for or nature of support. The dialogue between Pam Lomax and John Cowan vividly captures the qualitative form of support which is necessary for someone embarking on changing their practice. Given the potential defensiveness and vulnerability of those who expose aspects of their practice as areas for change, this is paramount. The two contributors write:

Pamela: "How do you cope with the fact that people will claim they hold certain values and then deny those values in their practices? I think it is a crucial issue when one is getting students to make contracts about their own learning."

John: "If I am being totally genuine, my immediate reaction is that this person who holds values and then denies them in practice is probably me. But I doubt if you could shift me simply by telling me that I am a hypocrite. I am ashamed to admit that will probably push me into a position in which I will passionately defend my actions, and probably believe that I am being misunderstood or misjudged. You must find a way in which you hold up a mirror or play back a tape recorder (literally or metaphorically) or encourage me to test out my values in a variety of developing situations until I discover, privately and in a supporting environment, that there is an inconsistency of which I am ashamed and which I wish to change. That's not, maybe, the only way to do it, but it's the only way I know. It works for me when someone caringly facilitates me in that style."

The need for the involvement of caring professional colleagues is a tenet to which others also subscribe. Several authors point to the benefits of collaboration and teachers working together to share problems, experiences and understanding of professional practice.

Whilst this text does not address itself directly to educational psychology, there are nevertheless some lessons for the profession. Namely that the profession should produce some case studies where its members can demonstrate, in their claims to educational knowledge, how they have facilitated teachers to improve their practice. Such studies would be important contributions to the development of a psychology which was truly "educational'.

Joan Whitehead

McManus, M. (1989). *Troublesome Behaviour in the Classroom.* London: Routledge, pp. 171, £25.00 hbk., £9.95 pbk. ISBN 0-415-03910-X; 0-415-03911-8.

It must be stated at the outset that the author is not the famous wrestler of that name and is therefore enjoined to employ more eclectic and reasoned approaches to pupil indiscipline. The book is an unusual mixture of approaches but this certainly does not detract from its readability. For most of the text the writing projects distinctly human qualities with Mr McManus using a cocktail of practical advice resulting from hard-won experience, academic references and case examples often laced with good humour.

The author is deputy head of a secondary school and it is stated the book is "written by a teacher for other teachers" and I think this is exactly so. The teacher who gained some experience would find this book at an appropriate level, full of professional common-sense and reasoning in putting discipline problems into perspective.

The student and probationary teacher would be better equipped with that old favourite *Effective Classroom Management* by Laslett and Smith (1984) or *Effective Classroom Control* by Robertson (1989). The latter deals much more thoroughly with defining and establishing the teacher's authority and relationship with pupils than McManus manages to do.

This is not particularly a book for psychologists or practising academics except where the author describes how the book might be used in an INSET capacity. Here, the indication is that such courses should be school-specific, run by teachers for teachers. He believes in teachers "consuming their own smoke" and gaining more self-respect by doing so. As far as (educational) psychologists go it is worth quoting in full:

"...The study of psychology and two or three years' teaching experience is no substitute for twenty years of reflective practice in the classroom.... The real experts to turn to are our colleagues...." (p. 155).

The main thrust of the book is in looking at common pupil discipline concerns (McManus studiously avoids calling them problems) from both the pupil and teachers' perspective and looks at the needs of both parties. That current 'Paracetamol', the whole school approach, is not particularly embraced in this book and although the book is recent there is no direct reference to the Elton Report and its useful recommendations.

Overall, the book is worth reading not least for the fresh and wholesome snippets sprinkled throughout the text, such as "teachers need to acquire increased confidence and changed attitudes to difficult behaviour: to see it as tractable, as something to be tackled dispassionately, as part of their job and not just as grit in the machine".

M. I. Smith

STEFFY, B. (1989). *Career Stages of Classroom Teachers.* Lancaster, PA: Technomic Publishing Co., pp. 213, n.p. ISBN 0-87762-673-5.

In-service training for teachers often fails to take account of teachers' different interests, working contexts and stages of professional development; it often focuses on technical skills; assumes the teacher is what needs to be developed rather than the school system; and is based on the expert talking to teachers but often failing to communicate what teachers want to know and be able to do. These are the criticisms that Steffy makes of current in-service provision for teachers in the United States, though many of these points might well sound familiar to teachers in other countries.

Following a strong and convincing critique of current professional development activity, Steffy draws upon a wealth of personal experience as a teacher and educational administrator to present and substantiate a model of teachers' careers that might guide the work of those who attempt to promote the professional development of teachers.

Steffy takes a predominantly humanistic stance, arguing that teachers need to be considered as individuals and that their personal and professional needs cannot be conveniently separated. The teacher who experiences personal anxieties, personality difficulties or human relationship problems requires help with these before any more direct attempt to change their classroom performance can be expected to take effect.

Five stages of teachers' careers are identified to aid the assessment of teachers' in-service needs: the anticipatory stage; the expert/master stage; the withdrawal stage; the renewal stage; and the exit career stage. Drawing upon various examples from her own experience, Steffy outlines the concerns and characteristics of teachers at these different career stages. The anticipatory stage, for instance, is characterised by idealism, energy, an openness to new ideas, creativity and an orientation to professional growth. The withdrawal stage, on the other hand, occurs at three different levels — initial withdrawal, persistent withdrawal and deep withdrawal — characterised by progressively greater cynicism, isolation, defensiveness and lack of enthusiasm.

Steffy suggests that teachers progress through these stages at different rates and in different ways. Typically, however, the anticipatory stage lasts for the first three or so years of a teacher's professional life, followed by several years at the expert/master stage and then, for many, by withdrawal.

Steffy's primary concern is for teachers to be helped to the expert/master stage and for the regeneration of this stage for teachers who are about to enter withdrawal. Successive cycles of the expert/master stage and a renewal stage are viewed as leading eventually to the stage of leaving the profession with feelings of satisfaction and accomplishment. Steffy estimates that only 5-10 per cent of teachers maintain their development at the expert/master level when, she claims, with appropriate school support, it ought to be possible for 70 per cent or more to be at this level. The emphasis in the book is therefore on strategies of supporting teachers, developing their confidence, helping them to grow professionally and to take responsibility for their own professional development. In this, Steffy takes quite an eclectic and pragmatic stance, suggesting a variety of strategies for school administrators, ranging from reward systems and the supervision of what teachers do, to the development of agreed purposes within the school and the improvement of human relationships. Steffy, however, does not underestimate the potential damage to children and the school of a teacher who has entered deep withdrawal, nor does she underestimate the difficulties that such teachers present for school administrators and the virtual impossibility of helping teachers to recover from this stage. Her advice is blunt and to the point: document their failings and urge them, both in their own interests and those of the school, to leave the profession!

Theoretically, the book draws upon Maslow's theory of motivation and upon Herzeberg's theory of work satisfaction, but much of what is presented draws essentially upon Steffy's own experience and reasoning about professional growth in teaching. From the research perspective, the account is weak in its coverage and critique of existing models of professional growth and work satisfaction, and in its lack of any systematically collected empirical evidence to support the proposed model. As with any stage model of career development, the dynamics of personal commitment, enthusiasm, effort and performance are somewhat simplified, and in particular the complex interactions between teachers' practice and children's learning is little explored. Nevertheless, for the practitioner, the book is rich in vignettes and examples of teachers' careers which have been well organised to illustrate strategies and their possible effects. For administrators and those concerned with supporting professional development in schools it represents an interesting, useful and very readable commentary on teacher development that could potentially provoke many discussions about what schools, teacher educators and policy-makers might do to improve the quality of work and satisfaction of teachers in school.

JAMES CALDERHEAD

WALKERDINE, V. *et al.* (1989). *Counting Girls Out.* London: Virago, pp. 224, £6.50 pbk. ISBN 1-85381-069-X.

This book aims to study a range of issues of interest to psychologists: cognition, affect, performance, anxiety and "confidence", in the context of mathematical learning. It aims to look critically at the notion of girls' underachievement in maths relative to boys from their early years to just before GCSE, whether there is any evidence for such underachievement, and, if not, why there is such a persistent belief in it. The book is a collective effort, based on an impressive body of data, produced by the Girls and Maths Unit at the London Institute of Education over about ten years.

Counting Girls Out is clearly committed to the understanding of, and the struggle against, female oppression, particularly in the learning of maths. At the same time, there is a great deal here that will be valuable for those with a general interest in child development and mathematics education.

The theoretical basis for this work is "post-structuralism", which can be seen as an attempt to synthesise several currents of social theory:

(i) structuralism, with its use of linguistics (a la Saussure) where meaning is produced by a "sign" as a relation between the "signifier" (typically a word or image) and the "signified";

(ii) the historical study of "discourses", how their significations change over time, while remaining linked to earlier meanings (see e.g., finding 1 below).

Walkerdine *et al.*'s approach also includes:

(iii) psychoanalytical ideas; and

(iv) the study of "lived experience" — how the teachers and the children understand and make sense of their activities — through intensive observation and interviews.

A quotation from the book will indicate the enormous range, interest and power of the approach:

"Mathematical meanings — indeed, the development of language and word meanings in general — cannot be separated from the practices in which the girls grow up. The mother is positioned as regulative in these practices, in which desires, fears, and fantasies are deeply involved. So 'mathematical meanings' are not simply intellectual, nor are they comprehensible outside the practices of their production. Yet in school ... children have to learn that there are special meanings to these terms, which are not necessarily those used at home. These meanings lead to the generation of mathematical statements of enormous power.... We analysed the transcripts of recordings of thirty mother-daughter pairs ... While there were many examples of 'more', 'less' did not occur once ... all instances of 'more' come from mother-daughter exchanges where the daughter's consumption of scarce or expensive resources and food is regulated by the mother.... The opposite of 'more' in food regulative practices is something like 'no more', 'not as much', and so on.... Here these terms, for the girls ... carry strong emotional ... content and act as signifiers in very different (ways) from the word pair 'more'/'less' as used in school mathematics.... Shifts from these practices and emotions to understanding 'mathematical terms' — in this case the 'more'/'less' pair in ... the comparison of quantities — have to be accomplished" (pp. 52-53).

Thus, different contexts, such as home and school, are characterised by — indeed, are constituted by — different practices and related sets of terms and meanings: "discursive practices". Subjects are put in "positions" by the practices in which they are engaged, as well as structurally, for example depending on gender; and also by a struggle for power in which the use of language is crucial. This is a lesson which children learn early, to judge from several transcripts provided by the authors, illustrating interaction among children (e.g., pp. 69-70, pp. 132 ff.), and one striking incident where two little boys, in resisting the teacher's requests, deploy terms from a sexist discourse in a way that threatens to challenge her control (pp. 65 ff.).

Here are some further representative findings:

(1) 19th century "discourses" that held women to be excessively swayed by emotions and therefore lacking in capacity for rational judgment live on in today's common sense that "women's

minds" are not fertile ground for maths and hard sciences. More specifically, in primary school maths, the discourses of teaching tend to view as laudable "active learning", "breaking set", etc. and to view as pathological "rote-learning" and rule-following. But these ideas are gendered: girls tend to be positioned in the social interactions of the classroom as neat, helpful, hard-working and well-behaved — and then their production of behaviour consistent with such "characteristics" tends to be read as evidence of their passivity, in contrast with boys' naughtiness and restlessness in the classroom — which is seen as testifying to their "potential", "mathematical flair", etc. Walkerdine and her colleagues here present evidence from interviews with both teachers and pupils.

(2) Teachers' methods and judgments position girls differently in maths and in English: the girls thus evaluate their abilities differently in the two subjects, which helps to explain their differing tendencies to persevere with the two subjects.

(3) If girls are positioned as "not successful", this is not because they are actually not succeeding in maths. Here Walkerdine and her fellow researchers examine the findings of national studies such as those of the Assessment of Performance Unit (APU), and present their own results of much smaller-scale testing in the schools where they observed. In the APU case, they point out that few gender differences are even statistically significant, and that even those would be equivalent to only two or three points on a test with standard deviation similar to the IQ. Their own results appear to show a similar lack of gender differences. Unfortunately the discussion for the 10-year-olds (pp. 118-120) is not presented clearly. For first year secondary, the boys excel on the test over girls, but the differences between tutor sets are as least as great as gender differences. For fourth year secondary, the girls are superior in all components of the mock examination (test, course work, investigations), again with large differences between tutor sets. The finding that gender differences claimed by some to exist in mathematics performance in secondary school cannot be replicated, even in a relatively small-scale study, merits attention and further research.

It is impossible in a short review to do justice to the breadth of the theoretical basis, or to the richness of the data produced. Yet I believe the approach illustrated here, and presented earlier in Walkerdine's *The Mastery of Reason* (1988) and Henriques *et al.*, *Changing the Subject: Psychology, Social Regulation and Subjectivity* (1984), is exceedingly fruitful both for research, and for understanding classrooms and the individuals in them. It shows how to study issues central to psychology, such as the links between cognition and affect, in a way that does not exclude equally important aspects of social life that are conventionally included under other social science disciplines — for example, social structure, material inequality, power (through the notion of "positioning"), and a concern with "lived experience".

The problems which strike this reviewer flow from the same breadth and richness:

—how to justify a particular historical interpretation of the meanings of a discourse, using small number of texts and small samples of quotations from within them (see the discussion (pp. 72-76) of the messages of the Plowden Report for an illustration of the difficulties); and

— how to justify psychoanalytical interpretations of, say, children's behaviour, based on relatively few interviews with the children, their parents and their teachers.

I do not wish to imply that these problems are insurmountable — there is substantial work on the first in semiotics, and scattered work on the second — but rather to the point to their importance in the development of the forward-looking work of which *Counting Girls Out* provides an exciting example.

JEFF EVANS

VALSINER, J. (Ed.) (1988). *Child Development within Culturally Structured Environments. Vols. 1 & 2.* Norwood, N.J.: pp. 260 and 305, £33.95 per volume, hbk. ISBN 0-89391-487-8; 0-89391-488-6.

These two volumes report earnest efforts to make theoretical progress in understanding early socialisation processes. Empirical work in this field has a long history which has until very recently been dominated by passive-receptive models of "culture". "Culture" was operationalised as "homogenous given". Subjects were "working class" or "American" or "black" for example. Families (generally embodied in "mother") transmitted through "rearing practices", "culture" to infants who absorbed it to greater or lesser degrees. It may be acknowledged that the rhetoric of this work often hinted at a richer, more diversified and more complex picture than this. The fact remains that empirical work adopted simple-undirectional and one dimensional models.

Several factors have conspired to render this work unsatisfactory. From within the field, whilst "findings" were amassed, little theoretical progress emerged. "Findings" were often contradictory and, in terms of the operations used, unintelligible. "Good mothers", for example, were "warm" and also "firm" said the correlational studies. But presumably not both at once. Basic questions were unanswered. Why did similar practices have such very different effects? Why within a cultural given, did practices vary so much? If cultures were "given" how did culture change? And so on. From outside the field the emergence of constructivist models of cognitive development posed challenges for the passive-receptive models of the child. A growing sophistication with correlational studies provoked questions about the direction of causation of "effects" reported in socialisation research. Did "warm" mothering cause "adjusted, calm" infants or did having a calm infant release "warm" mothering responses? More fundamentally, despite decades of research the fundamental question in the field, "what is the meaning of culture as a process in the ontogeny of children?" remain unanswered. It is the key question picked up in this collection. In the hands of the editor and his contributors it has become exceedingly — but very properly — complex.

They are to be congratulated on their sustained attempts to drive their work by explicitly articulating theoretical positions. These positions are not entirely similar, but they do carry a family resemblence. A "co-constructivist" model is adopted almost unanimously. This entails recognising that all participants in the socialisation process are in the business of constructing meaning from social interactions. In interpreting these interactions, the challenge for the psychologist is to understand the intersubjectivity of the participants. However, it is more complicated than that. Interactions take place in settings and with artefacts which are, to a degree, culturally determined, but not fixed. The artefacts and settings are part of the dynamics of the interaction. Nothing carries an immutable "sign". The psychologist, then, must, if socialisation is to be understood, grasp how settings are structured and operated on by the participants.

This general theoretical stance is fleshed out in different ways by different contributors. The broad range covered (from a study of the adjustment to "snack" time in US schools to a study of the role of cave paintings in initiation rites in pre-historic Europe) is beyond the scope of this reviewer so further comment is restricted.

Particularly striking is the effort by Ignjatovic-Savic and colleagues to develop the work of Vygotsky — and in particular the concept of "zone of proximal development". In recent experimental socio-cognitive psychology, which is almost entirely a-theoretical, Vygotsky's name has been used with neither critique nor explanation, as a kind of seal of approval. In the present volume, Vygotsky's ideas are put under serious scrutiny as a precursor to elaboration and operational definition before using them in a study of the developmental effects of social interaction. It is an open question as to whether the attempt is successful or not. What is noteworthy is that the attempt is made.

A second less auspicious feature of the collection is the lack of connection between the work reported here and current development in classroom research. This is especially important since classrooms are significant culturally structured environments. True, there is classroom research reported, but one piece is a study of social rituals in school, and the other, a comment on science teaching as indoctrination in the GDR, abandons entirely the theoretical frameworks set up in the other papers. More generally, the importance of social setting and of reciprocal interactions was recognised long ago in classroom research. More recently the dynamic influence of subject matter knowledge has been explored. In this domain, ecological approaches have shown some promise in explaining classroom life. Ecological models hardly feature in these volumes. A surprising omission.

Overall one gets a sense of little real progress in the field but, more promisingly, a sense of workers looking in profitable directions and revving up the engine of theory on their baggage cart. This is so refreshing that the tomes are worth serious study on these grounds alone.

CHARLES DESFORGES

PUBLICATIONS RECEIVED

Books for review have been selected from the following
("n.p." means "no price mentioned")

Anderson, E. (1989). *Speaking with Style*. London: Routledge, pp. 200, £30.00 hbk. ISBN 0-415-02256-8.

Apter, T. (1990). *Altered Loves*. Hemel Hempstead: Harvester Wheatsheaf, pp. 231. £25.00 hbk. ISBN 0-312-04534-4.

Barrett, G. (1989). *Disaffection from School: The Early Years*. Basingstoke: Falmer Press, pp. 177, £9.95 pbk. ISBN 1-85000-441-2.

Billig, M. (1989). *Arguing and Thinking*. Cambridge: Cambridge University Press, pp. 290, £9.95 pbk. ISBN 0-521-33987-1.

Bond, L., and Compas, B. (1989). *Primary Prevention and Promotion in the Schools*. London: Sage, pp. 452, £33.50 hbk. ISBN 0-8039-2908-0.

Bornstein, M. H., and Bruner, J. (1989). *Interaction in Human Development*. Hillsdale, N.J.: Lawrence Erlbaum Associates pp. 306, $49.95 hbk. ISBN 0-8058-0035-2.

Bowers, T. (1989) *Managing Special Needs*. Milton Keynes: Open University Press, pp. 206. £9.99 pbk. ISBN 0-335-09256-X.

Burgess, G. (1990). *The Song of Roland*. London: Penguin, pp. 219. £3.99 pbk. ISBN 0-14-044532-3.

Burman, E. (1989). *Feminists and Psychological Practice*. London: Sage, pp. 198. £9.95 pbk. ISBN 0-8039-8233-X.

Case, C., and Dalley, T. (1990). *Working with Children in Art Therapy*. London: Routledge, pp. 224, £14.99 pbk. ISBN 0-415-01738-6.

Clark, M. (1989), *Understanding Research in Early Education*. London: Harwood Academic Publishers, pp. 132, £9.50 pbk. ISBN 2-881-124-730-X.

Cole, T. (1989). *Apart or A Part*. Milton Keynes: Open University Press, pp. 189, n.p. pbk. ISBN 0-335-09225.

Coleman, J. C., and Hendry, L. (1989). *The Nature of Adolescence*. (2nd edit.) London: Routledge, pp. 255, £9.95 pbk. ISBN 0-415-01485-9.

Davies, D. (1989). *Psychological Factors in Competitive Sport*. Basingstoke: Falmer Press, pp. 177, £7.95 pbk. ISBN 1-85000-607-5.

Delgado-Gaitan, C. (1990). *Literacy for Empowerment*, Basingstoke: Falmer Press, pp. 182, £9.95 pbk. ISBN 1-85000-663-6.

Dillon, J. T. (1989). *The Practice of Questioning*. London: Routledge, pp. 287, £12.95 pbk. ISBN 0-415-04379-4.

Eisenberg, N., and Mussen, P. (1989). *The Roots of Prosocial Behaviour in Children*. Cambridge: Cambridge University Press, pp. 195, £8.95 pbk. ISBN 0-521-3371-2.

Eldering, L., and Kloprogge, J. (1989). *Different Cultures Same School*. Netherlands: Swets, pp. 262, n.p. hbk. ISBN 90-265-0989-8.

Eliot, G. (1989). *Middlemarch*. London: Penguin, pp. 161, £3.99 pbk. ISBN 0-14-077173-5.

Emery, R. (1988). *Marriage, Divorce and Children's Adjustment*. London: Sage, pp. 160, £12.00 pbk. ISBN 0-8039-2781-9.

Ernest, P. (1989). *Mathematics Teaching: The State of the Art*. Basingstoke: Falmer Press, pp. 279, £10.95 pbk. ISBN 1-85000-461-7.

Evans, R. (1989). *Special Educational Needs*. Oxford: Basil Blackwell, pp. 229, £22.50 hbk. ISBN 0-631-16902-4.

Furnham, A. (1990). *The Protestant Work Ethic*. London: Routledge, pp. 305, £12.99 pbk. ISBN 0-415-01705-X.

Green, R. (1989). *Human Aggression*. Milton Keynes: Open University Press, pp. 241, £9.99 pbk. ISBN 0-335-15345-3.

Harris, P. (1989). *Children and Emotion*. Oxford: Basil Blackwell, pp. 243, £9.95 pbk. ISBN 0-631-16753-6.

Holderness, G. (1989). *Richard III*. London: Penguin, pp. 105. £3.50 pbk. ISBN 0-14-077249-9.

Hollingdale, J. (1990). *Georg Christoph Lichtenberg: Aphorisms*. London: Penguin, pp. 202, £4.99 pbk. ISBN 0-14-044519-6.

Hollway, W. (1989). *Subjectivity and Method in Psychology*. London: Sage, pp. 150, £7.95 pbk. ISBN 0-8039-8208-9.

Kaye, K. (1982). *Mental and Social Life of Babies*. Hemel Hempstead: Prentice Hall, pp. 289, £40.00 hbk. ISBN 0-7108-4.

Kitwood, T. (1990). *Concern for Others*. London: Routledge, pp. 256, £12.99 pbk. ISBN 0-415-04377-8.

Kruglanski, A. (1989). *Lay Epistemics and Human Knowledge*. New York: Plenum, pp. 281, £35.00 hbk. ISBN 0-306-43078-9.

Lang, G., Van der Molb, H., Trower, P., and Look, R. (1990). *Personal Conversations*. London: Routledge, pp. 192, £9.95 pbk. ISBN 0-415-03478-7.

Lavern-Berry, G., and Keiko Asamen, J. (1990). *Black Students*. London: Sage, pp. 319, £13.95 pbk. ISBN 0-8039-2462-3.

Lighthall, F., and Allan, S. (1989). *Local Realities, Local Adaptations*. Basingstoke: Falmer Press, pp. 428, £13.95 pbk. ISBN 1-85000-588-5.

Newman, C. (1989). *Young Runaways*. London: The Children's Society, pp. 157, n.p. pbk. ISBN 0-907324-51-7.

Newman, D., Griffin, P., and Cole, M. (1989). *The Construction Zone*. Cambridge: Cambridge University Press, pp. 169, £8.95 pbk. ISBN 0-521-38942-9.

Radford, J., and Rose, D. (1989). *A Liberal Science*. Milton Keynes: Open University Press, pp. 146, £9.99 pbk. ISBN 0-335 09502-X.

Roaf, C., and Bines, H. (1989). *Needs, Rights and Opportunities*. Basingstoke: Falmer Press, pp. 153, £8.95 pbk. ISBN 1-85000-517-6.

Rose, M. (1990). *Healing Hurt Minds*. London: Routledge, pp. 210. £12.99 pbk. ISBN 0-415-01796-3.

Rowan, J. (1989). *Subpersonalities*. London: Routledge, pp. 242, £8.95 pbk. ISBN 0-415-04329-8.

Ryden, M. (1989). *Dyslexia: How Would I Cope?* London: Jessica Kingsley, pp. 64, £3.95 pbk. ISBN 1-85302-026-5.

Stainthorp, R. (1989). *Practical Psychology for Primary Teachers*. Basingstoke: Falmer Press, pp. 215, £8.95 pbk. ISBN 1-85000-562-1.

Starratt, R. (1989). *The Drama of Schooling/The Schooling of Drama*. Basingstoke: Falmer Press, pp. 151, £8.95 pbk. ISBN 1-85000-705-5.

Trueba, H., Spindler, G., and Spindler, L. (1989). *What Do Anthropologists Have to Say About Dropping Out?* Basingstoke: Falmer Press, pp. 155, £8.95 pbk. ISBN 1-85000-621-0.

Winter, R. (1989). *Learning from Experience*. Basingstoke: Falmer Press, pp. 201, £9.95 pbk. ISBN 1-85000-611-3.

Worell, J., and Danner, F. (1989). *The Adolescent as Decision-Maker*, New York: Academic Press, pp. 320, $49.95 hbk. ISBN 0-12-764052-5.

Wray, D., Bloom, W., and Hall, N. (1989). *Literacy and Learning in Action*. Basingstoke: Falmer Press, pp. 179, £9.95 pbk. ISBN 1-85000-605-9.

British Journal of Developmental Psychology

Special issue on 'The child's theory of mind'

Submissions are invited for a special issue of the *British Journal of Developmental Psychology* on 'The child's theory of mind'. Empirical and theoretical contributions concerning the development of the child's knowledge of her own mind and the minds of other people are requested. The topic may be approached from any developmental or comparative perspective and may include aspects of emotional development or the developmental psychopathology of metacognition. Papers not chosen for the special issue will be considered for the journal in the usual way. The issue will be published as part of volume 9 (1991).

Papers should be submitted to the Editor, Professor G. E. Butterworth, Journals Office, 13A Church Lane, London N2 8DX, UK *before 28 February 1990.*

The *British Journal of Developmental Psychology* is published four times a year. Subscription price of volume 7 (1989) is £55.00 (US$103).

A special individual price of US$54.00 is available to members of APA/SRCD/ISSBD.

The British Psychological Society
The Distribution Centre, Blackhorse Road, Letchworth,
Herts SG6 1HN, UK.

Journal of Occupational Psychology

Selected contents of recent issues

Rewards investments alternatives and organizational commitment: Empirical evidence and theoretical development — *N. Oliver*
Performance evaluation and compensation feedback messages: An integrated model — *B.D. Bannister & D.B. Balkin*
Coping with job-related stress: The case of teachers — *I.S. Schonfeld*
Self-ratings, supervisor ratings, position and performance — *P. Herriot & J. Lane*
Predicting commitment to company and union: Divergent models — *J. Barling, B. Wade & C. Fullagar*
Consequences of changing employers amongst young engineers — *T.J. Newton & A. Keenan*
Decision making in the graduate selection interview: A field study — *N. Anderson & V. Shackleton*
The measurement and antecedents of affective, continuance and normative commitment to the organization — *N.J. Allen & J.P. Meyer*
Union commitment: An examination of antecedent and outcome factors — *J.W. Thacker, M.W. Fields & L.A. Barclay*

The *Journal of Occupational Psychology* (ISSN 0305-8107) is published four times a year and edited by Dr Michael West (University of Sheffield). The price of volume 63(1990) is £64 (US$119.00).

Members and Foreign Affiliates of the Society and individual members of APA and other selected societies are entitled to purchase journals at preferential rates.

Subscriptions and inquiries to
The British Psychological Society
The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN

Journals of The British Psychological Society

British Journal of Psychology
Edited by A.J. Chapman, University of Leeds
The Journal publishes reports of empirical studies likely to further our understanding of psychology; critical reviews of the literature; and theoretical contributions. Readership is world-wide and includes academic and research workers in experimental and cognitive psychology.
Volume 81 (1990) (4 parts) £97.00 (US$181.00)

British Journal of Medical Psychology
Edited by John Birtchnell, Institute of Psychiatry
The British Journal of Medical Psychology publishes original contributions to knowledge in the area of those aspects of psychology applicable to medicine and related clinical disciplines, including the psychotherapies and counselling. Published articles include empirical studies, theoretical papers and clinical reports which suggest increased understanding of the psychological aspects of medicine.

The readers are medical psychologists, mainly clinicians, concerned with applying psychology to medical problems, from the disciplines of psychology, medicine and professions allied to medicine.
Volume 63 (1990) (4 parts) £76.00 (US$141.00)

British Journal of Social Psychology
Edited by Miles Hewstone, University of Bristol
The British Journal of Social Psychology publishes articles which contribute to the basic methodological and theoretical issues confronting the discipline. In addition, it encourages: theoretical and review papers; papers describing applied social psychology in a variety of settings; correspondence on controversial issues, guest-edited issues on research in growth areas of social psychology
Volume 29 (1990) 4 parts £65.00 (US$121.00)

British Journal of Developmental Psychology
Edited by George Butterworth, Professor of Psychology, University of Stirling
This new journal, launched in 1983, publishes authoritative scientific articles describing all aspects of development from infancy to old age, including educational implications and the problems of abnormal development.
Volume 8 (1990) (4 parts) £61.00 (US$112.00)

British Journal of Mathematical and Statistical Psychology
Edited by A.D. Lovie, University of Liverpool
The Journal publishes articles relating to any areas of psychology which have a greater mathematical or statistical or other formal aspect to their argument than is usually acceptable to other journals. Articles with a clear reference to substantive psychological issues are preferred, e.g. new models for psychological processes, new approaches to existing data, critiques of existing models and improved algorithms for estimating the parameters of a model.
Volume 43 (1990) (2 parts) £64.00 (US$119.00)

Journal of Occupational Psychology
Edited by Michael West, University of Sheffield
An international journal of industrial and organizational psychology which describes and interprets new research into people at work. Areas covered include personnel and vocational psychology; occupational social psychology; human factors and engineering and computer psychology; organizational psychology; behavioural aspects of industrial relations. Innovative or interdisciplinary approaches to problems. are particularly prominent.
Volume 63 (1990) (4 parts) £64.00 (US$119.00)

British Journal of Clinical Psychology
Edited by Ray Hodgson, Department of Clinical Psychology, Whitchurch Hospital, Cardiff
The British Journal of Clinical Psychology publishes new findings, theoretical, methodological and review papers bearing on the whole field of clinical psychology and includes descriptive and aetiological studies of psychopathology; applications of psychology to medicine and health care; studies of the assessment and treatment of psychological disorders; social and organizational aspects of psychological disorders and ill-health. Contributions combine scientific excellence with relevance to the concerns of practising clinicians, bridging the gap between academic and professional interests.
Volume 29 (1990) (4 parts) £65.00 (US$121.00)

All seven journals ordered at the same time for one address £433 (US$805)

The British Psychological Society

The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK.

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

British Journal of Psychology
British Journal of Medical Psychology
Journal of Occupational Psychology
British Journal of Social Psychology
British Journal of Clinical Psychology
British Journal of Developmental Psychology
British Journal of Mathematical and Statistical Psychology
The British Journal of Educational Psychology

The *British Journal of Psychology*, the *British Journal of Medical Psychology*, *Journal of Occupational Psychology*, the *British Journal of Developmental Psychology*, the *British Journal of Social Psychology* and the *British Journal of Clinical Psychology* appear in quarterly parts. The 1990 subscription prices, payable in advance, are £97.00 (US$181.00) for the *British Journal of Psychology*; £65.00 (US$121.00) for the *British Journal of Social Psychology*; £65.00 (US$121.00) for the *British Journal of Clinical Psychology*; £61.00 (US$112.00) for the *British Journal of Developmental Psychology*; £76.00 (US$141.00) for the *British Journal of Medical Psychology* and £64.00 (US$119.00) for the *Journal of Occupational Psychology*.

The *British Journal of Mathematical and Statistical Psychology* is issued twice a year, in May and November. The subscription rate, payable in advance, is £64.00 per volume post free (US$119.00 overseas).

Subscriptions to any of these journals may be sent to a bookseller or subscription agent, or to The British Psychological Society, The Distribution Centre, Blackhorse Road, Letchworth, Herts., SG6 1HN, UK. A special price of £433.00 (US$805.00) for all seven journals going to the same address is offered for 1990.

The *British Journal of Educational Psychology* is issued three times a year, in February, June and November. The prepaid subscription is £30.00 a year post free, or £10.50 each number. The prepaid subscription price in North America is $60.00 US. Journals are sent accelerated surface post to the USA. Subscriptions should be sent to the Scottish Academic Press, 139 Leith Walk, Edinburgh, EH6 8NS.

MANUSCRIPTS

For the British Journal of Psychology should be sent to Professor A. J. Chapman, Journals Office, 13A Church Lane, London, N2 8DX.

For the British Journal of Medical Psychology should be sent to Dr John Birtchnell, Journals Office, 13A Church Lane, London N2 8DX.

For the Journal of Occupational Psychology should be sent to Dr Michael West, Journals Office, 13A Church Lane, London, N2 8DX

For the British Journal of Social Psychology should be sent to Dr Miles Hewstone, Journals Office, 13A Church Lane, London, N2 8DX.

For the British Journal of Clinical Psychology should be sent to Professor R. Hodgson, Journals Office, 13A Church Lane, London, N2 8DX

For the British Journal of Mathematical and Statistical Psychology should be sent to Dr A. D. Lovie, Journals Office, 13A Church Lane, London, N2 8DX.

For the British Journal of Developmental Psychology should be sent to Professor George Butterworth, Journals Office, 13A Church Lane, London, N2 8DX.

For The British Journal of Educational Psychology should be sent in duplicate to Professor Charles Desforges, University of Exeter, St Luke's, Heavitree Road, Exeter, EX1 2LU.

Contributors are advised to consult the Style Guide issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price £3.50 post free, from The British Psychological Society, St Andrew's House, 48 Princes Road East, Leicester, LE1 7DR

All manuscripts should be submitted in duplicate.

OFFPRINTS

Contributors on application receive fifty copies of their papers free. Additional copies may be obtained in units of fifty copies.

REPRINTS

The Journal is available in reprint from Swets and Zeitlinger B.V., P.O. Box 810, 2160 SZ Lisse, The Netherlands; Vols. 21-29, 31-47.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £40.00.

They are entitled to purchase any or all of the Society's Journals at the special rates of £8.00 per Journal per annum

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester, LE1 7DR.

BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY,

VOL. 60, PART 2, 1990

CONTENTS

ARTICLES:

SKAALVIK, E. M., AND SLETTA, O. — Experiences in School as Determinants of Behaviour in a Discussion Situation 133

CHAPMAN, J. W., LAMBOURNE, R., and SILVA, P. A. — Some Antecedents of Academic Self-Concept: A Longitudinal Study 142

TOWNSEND, M. A. R., MOORE, D. W., TUCK, B. F., and WILTON, K. M. — Headings Within Multiple-Choice Tests as Facilitators of Test Performance 153

KULHAVY, R.W., STOCK, W. A. THORNTON, N. E., WINSTON, K. S., and BEHRENS, J. T. — Response Feedback, Certitude and Learning from Text 161

HIBBETT, A., and FOGELMAN, K. — Future Lives of Truants: Family Formation and Health-Related Behaviour 171

CHADWICK, D., YULE, W., and ANDERSON, R. — The Examination Attainments of Secondary Pupils who Abuse Solvents 180

BANKS, J., GRAY, C., and FYFE, R., — The Written Recall of Printed Stories by Severely Deaf Children 192

RESEARCH NOTES:

CAIRNS, E. — The Relationship Between Adolescent Perceived Self-Competence and Attendance at Single-Sex Secondary School 207

JOHNSTON, R. S., ANDERSON, M., PERRETT, D. I., and HOLLIGAN, C. — Perceptual Dysfunction in Poor Readers: Evidence for Visual and Auditory Segmentation Problems in a Sub-group of Poor Readers 212

BORG, M. G., and FALZON, J. M. — Teachers' Perception of Primary Schoolchildren's Undesirable Behaviours: The Effects of Teaching Experience, Pupil's Age, Sex and Ability Stream 220

BOOK REVIEWS 227

PUBLICATIONS RECEIVED 240

UK ISSN 0007-0998

Edu

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

NOVEMBER 1990

VOL. 60, PART 3

THE
BRITISH JOURNAL
OF
EDUCATIONAL PSYCHOLOGY

Edited by

Professor CHARLES DESFORGES

Assistant Editors:

Dr H. C. M. CARROLL

Dr D. FONTANA

Mr G. HARVARD

Dr P. TOMLINSON

Dr M. YOUNGMAN

Sub-Editor:

Dr S. CUNNINGHAM

Book Review Editor:

FLORA J. MACLEOD

Advised by the Editorial Committee:

Professor G. BROWN

Dr A. MANSTEAD

Professor P. W. ROBINSON

Dr K. WHELDALL

Vol. 60, Part 3

The *British Journal of Educational Psychology* is issued by the British Psychological Society once a term, in February, June and November. The subscription is £30.00 a year ($60.00 North America), or £10.50 each number.

Subscriptions and Orders should be sent to Scottish Academic Press, 139 Leith Walk, Edinburgh, EH6 8NS. Members of the British Psychological Society receive the *Journal* on special terms.

Papers for publication in duplicate should be sent to Professor Charles Desforges, University of Exeter, St Luke's, Heavitree Road, Exeter, EX1 2LU. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for The *British Journal of Educational Psychology* should be sent to: Flora Macleod, University of Exeter, St Luke's, Heavitree Road, Exeter, EX1 2LU.

Advertisements for this *Journal* should be sent to Scottish Academic Press.

Published by
SCOTTISH ACADEMIC PRESS, 139 Leith Walk, Edinburgh, EH6 8NS,
for the British Journal of Educational Psychology, Ltd.

All rights reserved

© 1990. The British Journal of Educational Psychology

PRINTED IN GREAT BRITAIN BY
BELL AND BAIN LTD., GLASGOW

Br. J. educ. Psychol., 60, 245-254, 1990

PROBLEM-SOLVING, ACADEMIC PERFORMANCE AND SITUATED REASONING. A STUDY OF JOINT COGNITIVE ACTIVITY IN THE FORMAL SETTING

By ROGER SÄLJÖ AND JAN WYNDHAMN
(Department of Communication Studies, Linköping University, Sweden)

SUMMARY. The article reports a study of group work in the school context aimed at solving the everyday problem of establishing the postage rate for a letter using a letter-scale and a postage table. The participants, pupils aged 12-13 years in the sixth form of the Swedish comprehensive school, co-operated in groups of three that were homogenised with respect to academic ability. The results reveal that all groups draw on school knowledge when solving the task, but that this does not necessarily lead to success in finding the correct postage rate. Furthermore, the higher the academic ability of students, the longer and more roundabout the problem-solving procedures are. Inappropriate solutions and procedures for handling the table appear just as frequently in all groups as suggestions at some stage of the work, but they are generally rejected during the longer and more elaborate interpretative work among the pupils of high academic ability. It is argued that the relationship between knowledge acquired in the formal setting and the handling of this everyday problem is diffuse. Successful solution of the problem cannot be described as the application of a standardised algorithm but seems more to rely on extensive interpretative work involving the comparison of alternative models of reasoning.

INTRODUCTION

A frequent critique of education, reappearing at regular intervals and receiving support from laymen as well as informed experts, focuses on the difficulties in providing students with the skills and problem-solving capacities that are useful in what is referred to as "real life". The tacit epistemology underlying such critique is the assumption that problems appear in recurrent patterns in different settings and that they can be reproduced in identical forms within the context of a formal institution. Practising on prototypical tasks in school, then, will result in individuals being able to deal with similar problems in other settings.

At a general and normative level, there is, of course, no reason to object to this construction of the role of the educational system in cultivating human intellectual competence. It is clearly the role and responsibility of formal schooling to contribute to the cognitive skills of students and to provide them with the intellectual tools that are important in the outside world. From the research perspective, however, this particular understanding of the relationship between one social context — or "system of activity" to use the language suggested by Leontiev (1981) — and others, seems to owe more to tradition and inherited assumptions than to in-depth analyses of how people solve problems in various settings (Donaldson, 1978; Scribner, 1984). A salient feature of these inherited perspectives is the Cartesian legacy of conceptualising human thinking in abstract terms as an autonomous domain and as essentially context-free (Rommetveit, 1988). In this view, mastery of a set of general principles and definitions should result more or less automatically in individuals being able to solve the problems to which these principles apply irrespective of context.

In recent years, several researchers have challenged the tenability of the assumed transfer of cognitive skills between situations (cf., for example, Rogoff, 1984; Brown *et al.*, 1989). This is not to say that there has not been research evidence around for quite some time indicating the difficulties people have in dealing with conceptually similar problems in different situations (see, e.g., Wason and Johnson-Laird, 1972; Gick and Holyoak, 1980). Even within the specific setting of schooling, Newman *et al.* (1984) show how very few students identify "problem isomorphs" when dealing with two permutation problems even though they have been intentionally designed to share a common, logical structure. With particular reference to the problems relating to the difference between formal schooling and "everyday" life, Carraher *et al.* (1985), in studying the activities of children street-vendors in Brazil, found that the children had considerable difficulties in the paper and pencil world of the school in handling the types of problems they could be seen mastering with perfection within the everyday activity of selling. Lave *et al.* (1984), in a similar vein, found that people were much more successful in solving quantitative problems when carrying out the everyday activity of shopping than they were when meeting identical problems as decontextualised tasks on paper (cf. Lave, 1988).

At a more general level, the assumption that the skills acquired in one setting will be applicable to and readily available for use in other settings can be understood as a reflection of what Schön (1983, 1987) refers to as "the positivist epistemology of practice". According to this epistemology, teaching and learning in formal institutions should be geared towards providing students with generalised competencies that can be applied to the appropriate problems in any situation. The underlying assumption is that we "learn" in school and then later "apply" our knowledge and skills in everyday life and work. This view can be conceived both as part of a pervasive theory of learning (learning goes on inside rather than outside the formal setting) and simultaneously it provides a view of the proper role of a social institution, the school.

Adopting a Vygotskian perspective, it is, however, obvious that this interpretation of the relationship between what goes on in different institutionalised realities rests on a narrow conception of thought and action that overlooks the social nature of human cognition, and the variation in motives, commitments and premisses for human activities that exist between various settings and that are essential to how we deal with problems (cf. Marková, 1982).

The aim of the present study is to scrutinise problem-solving activities by pupils under the specific premisses for thinking and action assumed by them to be valid in the particular setting of the school. In other words, the attempt is one of shedding light on problem-solving within an institutionalised context in which actors have learned to identify and adapt to certain practices and patterns of rationality (cf. Elbers, 1986). In saying that the context of school is institutionalised, it is not claimed that a clear line of division should be drawn between such settings and "everyday life". On the contrary, the activities associated with schooling form a prominent part of everyday life of young people. The point emphasised is of an epistemological nature with, we argue, theoretical and methodological consequences; all human thinking (and acting) takes place in a setting which is perceived as an integral background for the focal task that is to be attended to. The problem of what competencies transfer between settings remains largely unsolved.

A primary interest of the present study has also been to study problem-solving of a more concrete, everyday — rather than specifically academic — kind, in order to scrutinise how students, judged by the educational system to differ in academic ability, handle the problem of reading a table and what types of knowledge they draw on when attempting to do so.

METHOD

The participants were required to collaborate in small groups to establish the postage rate for a letter using a letter-scale and a postage table. This table was a copy of the official Postage Table of the Swedish Post which is a small folder distributed widely in society. The critical part of this table can be seen in Table 1

The letter provided for the problem-solving task weighed 120 grammes. This had to be established by the pupils by using the letter-scale provided. In this sense, the problem was practical in nature. The task was solved during a mathematics lesson and all interaction between pupils was audio-recorded. One of the authors (JW) was present during the sessions.

TABLE 1
EXCERPT FROM POSTAGE TABLE

Letters
Domestic
Regular letters (and picture postcards)

Weight (not exceeding grammes)	Postage
20	2.10
100	4.00
250	7.50
500	11.50
1000	15.50

Note that Swedish currency (SEK) is divided into kronor (crowns) and öre (100 to the crown)

Participants

In all, 45 pupils, aged 12-13 years, in the sixth form of the Swedish comprehensive school took part in the study. The task was solved collectively in groups of three pupils. The groups were formed so as to be homogeneous with respect to the level of school achievement. Since no grades are given at this age level, the pupils were placed into one of three achievement levels — low, average and high respectively — on the basis of their performance in mathematics. The assignment of the pupils to one of these categories was done by the regular class teacher. This means that the data derive from the problem-solving activities of 15 groups, five from each achievement level.

RESULTS

Generally speaking, the problem of establishing the postage rate turned out to be quite a challenge for the participants and only a third of the groups eventually ended up with the correct postage rate. It was also obvious that the students drew on knowledge acquired in school for establishing the postage rate, but that this in no immediate way seems to be linked to finding the correct answer.

There was a considerable variation in the approach taken to solve the problem and the time allotted to the task. One group used only 40 seconds to complete the entire task and agree on an answer (which was incorrect). The group using the longest time spent 23 minutes on the task (also arriving at an incorrect answer). Table 2 shows the time devoted to the task and the number of turns taken to reach an agreement in the group. A turn has been defined as the utterance(s) made by one individual within the same "linguistic move", i.e., until the conversation is taken over by the interlocutor. So called "backchannel talk" (short confirmatory utterances) has not been counted as turns.

TABLE 2

Time and Verbal Activity Devoted to Problem-Solving

Achievement level of group	Time devoted to task (secs)			Turns		
	Mean	Median	Range	Mean	Median	Range
Low	73	50	40-125	9.6	8	3-14
Average	433	165	60-1380	23.6	12	7-61
High	609	435	70-1080	35.6	29	11-66

The outcome pattern is consistent in the sense that the higher the achievement level in mathematics of the group, the longer time the group used and the longer the discussion took to come up with an answer that could be agreed upon. This perhaps somewhat contra-intuitive result hides, as we shall see below, distinctively differing approaches to solving the problem of establishing the postage for a letter. In statistical terms, the differences between groups in time used are significant (Kruskal-Wallis, $P<0.049$), while the differences in turns are not ($P=0.08$). The paired comparisons between groups can be seen from Table 3.

TABLE 3

Statistical Analyses of Differences in Time Used and Number of Turns for Solving Task (Mann-Whitney)

	Group	Turns		
		Low	Average	High
	Low		NS	0.016
Time	Average	0.016		NS
	High	0.016	NS	

Suggested postages rates and strategies for finding them

The first phase of the problem-solving activities in a situation such as this involves coming up with suggestions that can serve as raw material for the discussion on the postage rate. Counting all the postage rates that appear in the sessions, the groups produced a total of 37 suggestions. The most common one was 6.10 crowns, which was arrived at by *adding* the two postage rates for letters weighing 20 grammes (2.10 crowns) and 100 grammes (4 crowns) respectively. This was the suggestion made in S 322 and which was supported in subsequent turns by the other members. (Note that in the three-digit identification of each utterance, the first two digits refer to the achievement level of the particular group (11, 21, 31, 41, 51 = high achievers; 12, 22, etc. = average achievers; 13, 23, etc. = low achievers). The third digit refers to the individual student). The pertinent discussion in group 32 (average achievers) was as follows:

S 321: 120 grammes

S 322: 100 + 20 (grammes)

S 323: 6.10 (crowns)

S 322: 6.10 (crowns), is that right?

S 323: Let's agree on 6.10 (crowns)

S 322: (Adding) 4 (crowns) + 2.10 (crowns)

S 321: Yes

In 11 of the 15 groups, this postage rate was suggested and discussed, and in four cases this was the one agreed upon as correct. Two postage rates — 7.50 and 4.00 crowns — also appeared quite frequently. They were mentioned six times each. In both these cases, the postage rate was arrived at by using a different strategy of eliciting information from the table, namely by *reading* the table directly without introducing any computations. Group 23 arrived at the latter answer as can be seen from the following verbal exchange:

S 231: Should we weigh it then?

S 232: 120

S 233: Have you gotten anywhere?

S 232: 4 crowns

S 233: Me too

S 231: 4 crowns

S 232: 120 is between 100 and 250. It's closer to 100 than to 250. It's further from 250.

S 233: Answer: 4 crowns

The decision to agree on 4 crowns was based on a strategy which involved reading the table to find the appropriate interval and then rounding off to the closest value appearing, 100 grammes. This principle of rounding off corresponds to the mathematical principle acquired in school.

The postage rate of 7.50 crowns was arrived at by locating the interval and then deciding on the upper limit as was done by group 21 in the excerpt below. This, of course, is the correct alternative, since the table specifies the maximum weights in each interval.

S 211: (Weighing) 120 grammes

S 212: 7.50 (crowns) it should be

S 211: I think so too

S 213: Hmm

S 212: 7.50 (crowns) it is

S 211: Otherwise it should . . . Maybe one should put on two stamps?

S 212: No, we are not!

S 211: One can get fooled too (looks at interviewer)

S 212: It weighs *a maximum of* . . . (reading table, pointing at heading)

S 213: 7.50 (crowns) it is

S 211: Yes

A few other types of logic resulting in different answers should be added to this presentation. They represent attempts to arrive at a postage rate by interpolating into the table or by using the figures provided in accordance with a computational principle. In the following excerpt, group 51 suggested that the postage rate was 4.80 crowns by establishing that a letter weighing 100 grammes costs 4 crowns and then adding a fifth to make up for the additional 20 grammes. Thus, the problem was solved by means of a formal, computational approach.

S 511: Wait! (writes: 100 grammes — 4 crowns 1/5 of 100 grammes — 20 grammes, 1/5 of 4 crowns — 80 öre)

S 512: 4.80 (crowns) it is

S 511: Hmm

S 512: If you're to do it exactly, that's it. Otherwise one just takes . . . it'll be 7.50 (crowns)

S 513: Yes, 4.80 (crowns) is the exact amount.

An alternative way to establish the postage rate was to use a strictly proportional approach to calculate the postage per gramme and then to multiply by 120. Provided that one uses 4 crowns as the starting point for the calculation (rather than, for example, 2.10), the resulting postage is the same as in the previous excerpt.

S 412: Maybe we can take 4 crowns and divide by 100, then we'll know what 1 gramme costs

S 413: That'll be 4 öre

S 411: 4.80 (crowns) in that case

Yet another way of finding a principle for interpolating was exemplified by group 42. In the first turn by S 423, the pupil noted that 120 grammes is six times 20 which is the upper limit of the first interval.

S 423: 20 times 6 then? 2.10 times 6 then?

S 421: Should we calculate that? . . . 12.60 (crowns)

Table 4 gives a presentation of the outcome in terms of the total number of postage rates suggested in the groups.

TABLE 4

TOTAL NUMBER OF SUGGESTED POSTAGE RATES

	Suggested postage						
	4	4.80	6.10	7.50	12.60	Other	Total
Low	2		3	1		1	7
Average	2		4	1	1	6	14
High	2	2	4	4	1	3	16
Total	6	2	11	6	2	10	37

As is evident, the number of alternatives considered is higher among the average and high achievers than among the lower achievers (Mann-Whitney test for high vs. low achievers P=0.016, and for average vs. low achievers P=0.048). The former two groups introduce an average of 3.2 and 2.8 suggestions respectively, while the corresponding value of the latter group is 1.4. In one sense, this is an unexpected and to some extent contra-intuitive result: the higher the achievement level, the more tentative and roundabout the process of solving the problem appears and the more incorrect alternatives are suggested.

Deciding on a postage rate

The next step in the problem-solving procedure is one of deciding on a postage rate among the suggestions made, i.e., in those groups where more than one alternative has been considered. Analytically, this is a different phase of the groups' work. Here, criteria have to be considered and evaluated against assumptions about postage rates in an outside reality. In Table 5, the postage rates agreed upon by the members of the various groups as their answer is given.

TABLE 5

POSTAGE RATES DECIDED ON

Group	Low	Achievement level Average	High
1	6.10	4.20	7.50
2	4	7.50	7.50
3	4	6.10	7.50
4	6.10	No answer	4.80
5	6.10	4	7.50

Four of the groups of high achievers arrived at the correct postage rate, while only one of the groups of average and none of the low performers did so. At the statistical level, this means that the comparison between high and low achievers is significant (Fisher test, $P<0.025$), while the other differences are not. As can be seen, however, the incorrect solutions are by no means randomly distributed. The postage rates of 6.10 and 4 crowns were decided on in four and three groups respectively. This means that, with the exception of three groups, the outcome can be described in terms of three different answers.

A corollary of the finding that the majority of the answers can be reduced to a limited number of solutions is that, in terms of the semiotics of the reasoning underlying the problem-solving, there are a few, parallel interpretations of how to understand and handle the task. Furthermore, comparing the final solutions given by the groups with the total number considered, an interesting pattern emerges. The alternative of 4 crowns is mentioned in two groups at each achievement level and the postage rate of 6.10 is most frequent among the high and average achievers as a considered alternative but *not* as the final answer. The decisive difference between the groups, then, does not appear to be one of merely finding the correct alternative, but more one of judging what the relevant criteria are for choosing among different procedures of finding the postage rate. It is — we shall argue — the differences in the latter respect which appear crucial for explaining the variation in success these groups had in arriving at a correct answer.

Problem definition and problem-solving

Perhaps the most salient difference between groups has to do with the way in which the suggestions for solutions were handled. One obvious strategy — displayed in the excerpt from group 23 above — was to immediately accept a suggestion that was made and that was perceived as coherent according to some criterion. Consider also the following excerpt where the postage decided on was 6.10 crowns.

S 321: 120 grammes

S 322: 100 plus 20

S 323: 6.10 (crowns)

S 322: 6.10 Is that it?

S 323: We agree on 6.10 . . .?

S 321: Yes

The approach taken by groups 23 (above) and 32 was characterised by the fact that the solution of the problem was suggested more or less immediately without any consideration of alternative interpretations. There was, in other words, no explicit hermeneutic process involved where the interpretation of the problem was thematically attended to. A second decisive feature was that the arguments considered when attending to the problem took their point of departure in the suggestion made and provided proof of the correctness of the suggestion. Thus, the arguments offered all *supported* the solution considered.

The differences between the achievement groups do, however, not seem to result from a difference in the ability to arrive initially at a correct answer. This is evident from the more roundabout way in which the high achievers arrive at the answer, and also from an analysis of the initial solution suggested by the various groups. As can be seen from Table 6, there was a strong commonality between groups in the sense that there was a limited number of postage rates that appeared as first answers and that these were represented among all the groups.

TABLE 6

FIRST POSTAGE RATE SUGGESTED

Suggested postages	Low	Achievement level Average	High	Total
4	2	2	1	5
5	1		1	2
6.10	2	2	1	5
7.50		1	2	3

Instead, the main difference between groups, and the one that produces the more extensive discussions among the high achievers, is that the communicative activities are geared towards attempting to establish the proper definition of the problem by challenging suggestions made and providing alternatives. This interpretative work involves attending to a broad range of factors that are potentially relevant and that can account for the problems and the inconsistencies that the groups detect and that they have to settle during the problem-solving process. These problems concern arguing about the logic of the table and why the costs per unit decrease as the letter gets heavier, if it is reasonable to assume that the clerk at the post office will start adding or dividing to arrive at postages, if the distance you send a letter affects the postage, to mention just a few of the types of arguments that arise and that have to be resolved once they are brought into open discussion. This explicit discussion of how to interpret the problem, and what counts as valid arguments in this context, leads to more verbal activity and, thus, to a longer and more time-consuming problem-solving phase. It also introduces the very significant activities of *argumentation* and *context building* into the problem-solving, activities that cannot be found among the low achievers (Säljö and Wyndhamn, 1990).

DISCUSSION

The successful solution of the task of establishing the postage rate does not rely on the application of a standard algorithm that the participants bring to the situation and through which the problem becomes immediately transparent. Furthermore, all solutions — whether resulting in a correct or an incorrect postage rate — draw on knowledge acquired in school and in this sense even the mistakes made at various stages are founded on following the problem-solving procedures learned in the formal setting. This result also, in our view, lends support to the notion that the problem of applying school-based knowledge in specific settings is more complex than is conventionally assumed (cf. Schliemann, 1988). One immediate observation that can be made is the apparent unfamiliarity among the participants with tables of the kind used. Finding the correct postage rate in this study relies on extensive semiotic work and includes attempts at reconstructing the principles behind the table rather than on a direct reading of it. The problem experienced by students is not merely one of finding the appropriate postage rate, but more fundamentally one of deciding how to approach the task. The routes taken by several groups of adding postage rates, or, alternatively, making calculations as to what the per gramme postage is and then multiplying by 120, or interpolating into the table, violate the premisses for using the table, yet are approaches that are used by all groups.

A second observation, following this one, is that the handling of the problem clearly mirrors the fact that the task is being carried out in a school setting and during a mathematics lesson. Several of the modes of dealing with the problem can be read as attempts at reading some mathematics into it. The results also reflect the contextual knowledge demanded by students to realise that, for instance, dividing and multiplying or interpolation are not relevant strategies for solving the problem in spite of its appearing within a lesson allotted to mathematics. It is also obvious that with respect to the postage rate of 4 crowns, suggested by three groups, the answer is arrived at by strictly applying a principle of rounding off to the closest value which is mathematical in nature but which cannot be used in the case of the table. The setting, thus, is not external to the problem but poses a sense-giving background that is integral to solving it (Rogoff, 1984; cf. Säljö and Wyndhamn, 1987, 1988).

In our interpretation, what the results illustrate are the outcomes of fairly global differences between the groups when confronted with a task in the context of school, rather than in their specific abilities to handle that particular task as such. What differentiates groups are thus not primarily their solutions to the task in a technical sense, but rather their definitions of the situation and their willingness to engage in the type of interpretative work that the task demands under the specific institutional constraints typical of the formal setting. Let us attempt to clarify this.

Genuine problem-solving results from the experience of having an obstacle in the process of dealing with a situation. The problem-solving thus is a means towards what is perceived as a worthwhile end. Instances include handling a monetary transaction, mending a broken artefact or retrieving information. In such a situation, the nature of the problem is directly perceived, as is the range of possible means that might be instrumental in dealing with it. One could attempt to solve the problem on one's own, or, which is a common strategy, ask someone who is more knowledgeable. In the present situation of finding a postage rate, one could ask an adult for assistance, which is in fact a suggestion made by several participants. In the particular context of the school, usage of such sources of expertise is excluded by definition. The concrete problem-solving in this formal setting thus involves attending to a hypothetical problem, the solution of which is not an end in itself but a means to another goal, i.e., learning. At the same time there are clear restrictions on the resources that are appropriate for arriving at a solution where some potentially effective ones are excluded. The consequence of this institutional framing of the problem-solving process is thus that the students' intellectual activities are kept within specific boundaries as to the "structuring resources" (Lave, 1988) available for solving the task.

Furthermore, when looking at this task as an instance of a "school task", an interesting question is what competencies may be acquired by the participants in the various groups as a consequence of the work. To function as a learning task providing participants with a "semiotic challenge" (Wertsch, 1985), the pursuit of which is meant to contribute to their competence in handling similar tasks, the challenge has to be accepted. The differences between the groups in terms of the time and energy invested do seem more to concern the willingness to engage in the type of cognitive activity that is needed than to solve the factual task of establishing the postage rate in a realistic situation. For the groups characterised as high achievers the task seems to be accepted as an authentic one, mobilising considerable intellectual activity and concern in response. For the groups defined as low achievers the problem — in spite of its "realistic" character — does not seem to be interpreted as anything but a poorly disguised mathematical problem to be expected in these specific circumstances but which has no meaning in a wider existential sense.

In our view, the results illustrate how the particular definition of cognitive activities and learning prevalent in the school as a social institution feeds back on itself and serves as self-fulfilling prophecy. The institutional designation in this study of some groups as high achievers and other groups as low achievers cannot be considered as constituting an independent variable in a neutral sense depicting differences in skills of a purely cognitive nature. The success of problem-solving can — we venture to claim without direct support in our data — to a significant degree be understood as a function of the extent to which children feel concerned by the tasks given to them. In this perspective, in-depth scrutiny of what constitutes a learning task in the formal setting in itself will become a problematic and fascinating topic of research, as will the tradition of presenting students with tasks that do not mobilise their capacities in spite of the widely acclaimed necessity of promoting learning.

Acknowledgment.—The research reported here has been financed by the Swedish Council for Research in the Humanities and Social Sciences.

Correspondence and requests for reprints should be addressed to Dr R. Saljö, Department of Communication Studies, Linköping University, S-58183, Linköping, Sweden.

REFERENCES

Brown, J. S. Collins, A., and Duguld, P. (1989). Situated cognition and the culture of learning. *Educ. Researcher*, 18, 32-42.

Carraher, T. N., Carraher, D. W., and Schliemann, A. D. (1985). Mathematics in the streets and schools. *Br. J. dev. Psychol.*, 3, 21-29.

Donaldson, M. (1978). *Children's Minds*. Glasgow: Fontana.

Elbers, E. (1986). Interaction and instruction in the conservation experiment. *Eur. J. Psychol. Educ.*, 1, 77-89.

Gick, M. L., and Holyoak, K. J. (1980). Analogic problem solving. *Cogn. Psychol.* 12, 306-355.

Lave, J. (1988). *Cognition in Practice*. Cambridge: Cambridge University Press.

Lave, J., Murtaugh, M., and de la Rocha, O. (1984). The dialectic of arithmetic in grocery shopping. In Rogoff, B. and Lave, J. (Eds.), *Everyday Cognition: Its Development in Social Context*. Cambridge, Mass.: Harvard University Press.

Leontiev, A. (1981), The problem of activity in psychology. In Wertsch, J. V. (Ed.), *The Concept of Activity in Soviet Psychology*. New York: Sharpe.

Marková, I. (1982). *Paradigms, Thought and Language*. Chichester: Wiley.

Newman, D., Griffin, P., and Cole, M. (1984). Social constraints in laboratory and classroom tasks. In Rogoff, B. and Lave, J. (Eds.), *Everyday Cognition: Its Development in Social Context*. Cambridge, Mass.: Harvard University Press.

Rogoff, B. (1984). Introduction: thinking and learning in social context. In Rogoff, B. and Lave, J. (Eds.), *Everyday Cognition· Its Development in Social Context* (pp. 1-8). Cambridge, Mass.: Harvard University Press.

Rommetveit, R. (1988). Literacy and the myth of literal meaning. In Säljö, R. (Ed.), *The Written World*. New York: Springer.

Säljö, R., and Wyndhamn, J. (1987). The formal setting as a context for cognitive activities. An empirical study of arithmetic operations under conflicting premisses for communication. *Eur. J. Psychol. Educ.*, 2, 233-245.

Säljö, R., and Wyndhamn, J. (1988). Cognitive operations and educational framing of tasks. School as a context for arithmetic thought. *Scand. J. educ. Res.*, 32, 61-71.

Säljö, R., and Wyndhamn, J. (In press). Solving everyday problems in the formal setting. In Chaiklin, S. and Lave, J. (Eds.), *People in Action*. Cambridge, Mass.: Cambridge University Press.

Schliemann, A. D. (1988). Understanding the combinatorial system: development, school learning, and everyday experience. *Quarterly Newsletter of the Laboratory of Comparative Human Cognition*, 10, 3-7.

Schön, D. (1983). *The Reflective Practitioner. How Professionals Think in Action*. New York: Basic.

Schön, D. (1986). *Educating the Reflective Practitioner: Toward a New Design for Teaching and Learning in the Professions*. San Francisco: Jossey-Bass.

Scribner, S. (1984). Studying working intelligence. In Rogoff, B. and Lave, J. (Eds.), *Everyday Cognition: Its Development in Social Context*. Cambridge, Mass.: Harvard University Press.

Wason, P. C., and Johnson-Laird, P. N. (1972). *The Psychology of Reasoning: Structure and Content*. London: Batsford.

Wertsch, J. (1985). *Vygotsky and the Social Formation of Mind*. Cambridge, Mass.: Harvard University Press.

(*Manuscript received 12th November*, 1989)

Br. J. educ. Psychol., 60, 255-265, 1990

THE EFFECTS OF LIMITED PRIVATE REPRIMANDS AND INCREASED PRIVATE PRAISE ON CLASSROOM BEHAVIOUR IN FOUR BRITISH SECONDARY SCHOOL CLASSES

BY STEPHEN HOUGHTON, KEVIN WHELDALL, ROD JUKES
AND ANTHONY SHARPE
(Centre for Child Study, University of Birmingham)

SUMMARY. Four secondary school teachers were systematically observed teaching four different classes. Measures of class on-task behaviour and teacher use of praise and reprimand were made during each observation session. Following baseline observations, two teachers were asked first to decrease their use of reprimands and to make them more private and then to increase their use of private praise. The other two teachers were asked first to increase their use of private praise and then to add limited use of private reprimands. A multiple baseline design was employed so that one teacher in each pair introduced the intervention one week after the other teacher. Observations of teacher behaviour confirmed that all four teachers were able to comply with these instructions. The results showed clearly that both minimal use of private reprimands and use of private praise statements were effective in increasing the on-task behaviour of secondary aged pupils in all classes, by an average of over 20 per cent. Follow-up observations after two months on three classes showed that improved on-task behaviour was still apparent.

INTRODUCTION

NUMEROUS studies have now been completed which demonstrate the effective use of contingent praise as a classroom management strategy in British secondary school classrooms (Wheldall and Austin, 1980; Wheldall and Merrett, 1985, 1987; Scott *et al.*, 1986; Merrett and Wheldall, 1987). There has been little research, however, on the effectiveness of reprimands.

Wheldall *et al.* (1989) found that secondary teachers generally approved more often than they disapproved, giving on average up to three times more approval to pupil academic behaviour as disapproval. However, three times as much disapproval as approval was directed towards classroom social behaviour. Reprimands are clearly the most common form of punishment used by secondary school teachers. Just as teachers attempt to increase certain behaviours through the contingent use of approval, they also attempt to reduce other behaviours through the contingent use of disapproval. Van Houten *et al.* (1982) suggest that the paucity of experimental classroom based research on reprimands may be due to the inconclusive nature of earlier work with respect to the effectiveness of reprimands. They go on to suggest that the way in which reprimands are delivered and the varied interpretation of what constitutes a reprimand may be possible explanations for this.

O'Leary and O'Leary (1977) reviewed the studies carried out on reprimands in North America, which involved younger children. For example, Jones and Miller (1974) decreased high frequencies of disruptive behaviour in two special education classrooms by training the teachers to use reprimands in the same way as two more successful teachers. The less successful teachers were trained to reprimand misbehaviours briefly at their inception, in close proximity to the pupils, giving a clear explanation and accompanied by a facial expression and tone of voice consistent with disapproval. The teachers were also instructed to praise pupils for behaving appropriately following a reprimand. Jones and Miller further reported that the number of reprimands required to maintain order declined as the school year progressed. This suggests that reprimands can have long-term effects on performance if consistently applied.

In the light of Kazdin and Klock's (1973) finding that non-verbal behaviour and physical contact can increase the effectiveness of praise, Van Houten *et al.* (1982) examined the effects of reprimands on two children when delivered with and without eye contact and a firm grasp. Results indicated that although verbal reprimands delivered without contact were effective in both cases, the addition of eye contact and a firm grasp markedly increased their effectiveness. In a second study they compared reprimands delivered from within one metre of a pupil with reprimands delivered from seven metres. It was found that those delivered from within one metre were more effective, bringing about a greater reduction in disruptive behaviour. Only one person was involved in this experimental study, however. In a final study, the same authors examined whether reprimands delivered to one pupil affected the behaviour of an adjacent peer. This was found to be the case, with the amount of reduction in a given student's misbehaviour being essentially the same when his neighbour was reprimanded as when he was reprimanded himself.

A sample of two children from each of five grade two and three classrooms were selected by O'Leary *et al.* (1970) to participate in two studies comparing the effectiveness of loud and soft reprimands. Soft reprimands were seen by the authors as advantageous over loud reprimands since the former did not single out a pupil, thereby making his behaviour noticeable to others. Moreover, teachers perceived soft reprimands as a viable alternative to the usual methods of dealing with misbehaviour. Overall, the findings revealed that when teachers used soft reprimands they were effective in modifying behaviour in seven of the nine disruptive pupils. They concluded that, used in combination with praise for appropriate behaviour, soft reprimands may be helpful in reducing disruptive behaviour. In contrast, however, it appears that loud reprimands lead to a vicious circle of increasingly frequent reprimands, resulting in even more disruptive behaviour. Soft reprimands may also prove more reinforcing for the teacher, since a relatively small expenditure of effort often produces positive (and dramatic) changes in pupil behaviour.

O'Leary (1985) subsequently described a series of experimental studies on punishment conducted with young children. The studies fell into two main categories, namely parameters of reprimands and the interaction of punishment and reinforcement (praise) contingencies. She concluded that:

> "... timing and intensity are powerful variables, at least for some behaviours; the schedule or consistency of reprimands may be less important than we originally thought; length of reprimand is a weak but significant factor with brief reprimands being more effective; and voice tone may correlate meaningfully with other parameters but may not be a critical variable by itself." (p.9).

The findings of her three studies into the interactive effects of reprimands and praise showed that systematic removal of reprimands and praise (regardless of order) resulted in deterioration of pupil on-task behaviour when negative feedback was withdrawn but not when positive feedback was omitted. In a further study to address the generality of earlier findings, the teacher used only praise and reprimands. Praise did not improve pupil on-task behaviour achieved with reprimands and a deterioration in on-task behaviour was observed when reprimands were withdrawn. O'Leary believes that this demonstrates the importance of reprimands.

Hogan and Johnson (1985) argue that age appropriateness and perceived "fairness" of punishment may also be important variables when dealing with adolescents. A sharp, firm voice may elicit the desired response from a misbehaving young child, but may not have the same effect on a secondary pupil. Consideration should also be given to the possibility that reprimands and praise are mutually enhancing. Praise may be more effective when behaviour is occasionally reprimanded and reprimands more effective when behaviour is occasionally praised. It is known that punishment is most effective when paired with reinforcement for alternative behaviour (Azrin and Holz, 1966).

Pupil perception of praise is also an important consideration, especially in the secondary school context. Teacher praise may be highly embarrassing to sensitive adolescents especially when delivered publicly, in a loud voice. Houghton *et al.* (1988) have shown, in a recent survey of pupil opinion, that secondary pupils of all ages and both sexes prefer quiet, private praise to loud public praise for both academic achievement and classroom social behaviour.

In view of the lack of experimental research on reprimands in British secondary school classrooms the following study was carried out to examine the effects of private praise and private reprimands as classroom behaviour management strategies.

METHOD

Sample and setting

The experiment was carried out in a large inner-ring co-educational comprehensive school in the West Midlands. The study involved four secondary school teachers who had volunteered to take part in the study and their chosen respective classes. Teacher A (female) was observed teaching English to a third year class of 27 pupils (13 boys, 14 girls); teacher B (male) was observed teaching geography to a second year class of 25 pupils (12 boys, 13 girls); teacher C (male) was observed teaching mathematics to a fifth year class of 26 pupils (12 boys, 14 girls); whilst teacher D (male) was observed teaching chemistry to a fourth year class of 22 pupils (13 boys, 9 girls). The teachers had, on average, 13 years teaching experience (range nine to 18 years). All four classes comprised pupils of above average ability.

Observation schedule

The observation schedule employed was the OPTIC (Observing Pupils and Teachers In Classrooms) schedule developed by Merrett and Wheldall (1986). A copy of an earlier but substantially similar schedule (CTP2) was previously published in Wheldall *et al.* (1985), and was subsequently summarised in Merrett and Wheldall (1987). Consequently, only a brief description will be provided here. The development of OPTIC is described fully in Merrett and Wheldall (1986) together with evidence for its reliability and validity. Full instructions for its use are also provided.

Briefly, OPTIC is meant to be used when the whole class being observed is engaged in classroom activity either as a whole, in groups or individually. A complete observation takes 30 minutes. OPTIC allows the observer to sample two main aspects of classroom behaviour. In Section A the observer is concerned with positive and negative teacher responses to children's academic and social behaviours. Instructional language is ignored.

Section B is concerned with estimating children's on-task behaviour. The class is divided by eye into three groups of approximately equal numbers of pupils. The members of each group of pupils are then observed individually and in sequence for four seconds. If they are seen to be on-task for all of the four seconds then they are marked on-task. If, on the other hand, they are off-task at any time during the four seconds then they are recorded as off-task. Attention is then paid to the second group, then the pupils in the third group and so on. An estimate of overall percentage on-task behaviour for the class is subsequently calculated by summing the number of occasions when pupils were observed to be on-task and dividing by the total number of pupil observations.

The observer alternates between the two sections at three-minute intervals. Thus obtained rates of teacher behaviour relate to a total of a quarter hour observation time per session. In accordance with the instructions for using OPTIC each teacher and class was observed, in the present study, when the teacher and pupils were engaged in academic classroom-based activities. Classes were observed at different periods during the day and week in order to obtain representative samples of teacher and class behaviours. The data on each teacher and class were aggregated over the number of observation sessions conducted for each intervention phase to provide mean levels for each phase.

In the present study, teacher positive (praise) and negative (reprimand and disapproval) responses to pupils were further subdivided into categories denoting "public" and "private" responses. "Public" responses would be heard by the majority of the class, whereas "private" responses would be heard only by the target pupil and possibly by a peer sitting at the same table. Observations were carried out by the third and fourth authors who had been thoroughly trained in the use of OPTIC as part of a structured ten-week course in classroom behaviour management.

Prior to any observations taking place, the two observers visited each classroom on at least three separate occasions in order that the teachers and pupils would become accustomed to their presence in the classroom.

Experimental design

Baseline data were gathered on all four teachers over a period of two to three weeks comprising six separate observation sessions of both teachers A and B, and five separate observations of teachers C and D. Lessons occurred on only two or three days per week. Timetabling constraints dictated when the four classes could be observed, i.e., when they were being taught by that subject teacher.

Examination of the baseline data shows that the four teachers fell into two convenient pairs. Pair one comprised teachers A and B who both employed high rates of reprimand, whilst pair two comprised teachers C and D who employed low rates of feedback overall (i.e., positive and negative responding). These findings are described in more detail under "Results". This difference was incorporated into the multiple baseline experimental design.

The aim of the experiment was to evaluate the effects of the contingent use of quiet, private teacher praise and reprimands on pupil on-task behaviour. To this end, following baseline observations, contingent use of quiet reprimand was introduced into classes A and B whilst use of quiet praise was introduced into classes C and D. This was followed by the addition of quiet praise in classes A and B and the addition of quiet reprimands in classes C and D. In all cases, the interventions were introduced in classes A and C one week prior to their introduction in classes B and D, in accord with a minimal multiple baseline design (Kazdin, 1984). These interventions are detailed below under "procedure". A multiple baseline design was employed in preference to a reversal (return to baseline) design for both ethical and practical reasons. Teachers are reluctant to revert to previous procedures once significant improvements in pupil behaviour have been achieved as a result of changing their own behaviour. A multiple baseline design demonstrating change in the dependent variable when (and only when) the independent variable is manipulated obviates the need for reversals.

Procedure

Teachers A and B. Following six separate baseline observations the first intervention phase was introduced. As can be seen from Figure 1, teacher A began her intervention one week before teacher B. The intervention involved the use of quiet private reprimands. Teachers A and B were shown (separately) a video sequence on "Effective Reprimands" from the BATSAC package (Merrett and Wheldall, 1988). It was explained that to reprimand more effectively the teacher should move closer to the pupil, make eye contact and then reprimand briefly and privately, stating explicitly what is being disapproved of. These points were duplicated in the form of a written handout to act as a reminder. Both teachers were instructed to reduce their total reprimands to just a few quiet private ones. Teacher A was requested to reduce her number of (private) reprimands per double lesson (i.e., 70 minutes) to ten (2.1) or fewer (private) reprimands per double lesson. (In order to facilitate comparison with observed rates subsequently achieved, these rates are also given (in brackets) per quarter hour, i.e., the total teacher observation time per session.)

A further five separate observation sessions were made of each teacher over a period of three weeks and then the second intervention phase was introduced. Once again teacher A began this phase one week before teacher B. This intervention involved quiet private reprimands plus quiet private praise. Teacher A and B were shown (again separately) a video sequence on "Effective Praising", again from the BATSAC package. It was explained that to praise more effectively the teacher should move closer to a pupil, make eye contact and to praise briefly and privately, stating exactly what was being approved of. Private reprimand conditions were maintained and both teacher were asked to give at least 15 (3.2) approving (private) responses per double lesson. These points were again duplicated in the form of a written handout. Five further observations were made during this phase.

Teachers C and D. Following five separate baseline observation sessions teachers C and D followed a similar procedure to that of teachers A and B except that the treatments were reversed. The first intervention comprised a private praise phase. Teachers C and D were shown (separately) the video sequence "Effective Praising" and given the same instructions (and written handout) as teachers A and B received, except that at least 20 (4.2) approving (private) responses per double lesson were required. Five separate observation sessions were made of each teacher over a period of three weeks.

In the second, private praise and private reprimands phase, teachers C and D were shown (separately) the video sequence "Effective Reprimands" and given similar instructions (and written handout) that teachers A and B had previously received. They were told that they could give up to five (1.1) private reprimands per double lesson. Five separate observation sessions were made of each teacher over a period of three weeks. As can be seen from Figure 1 teacher C began each phase of intervention one week before teacher D.

The numbers for private praise and private reprimands requested from each teacher were realistic individual targets based on their relative baseline behaviour. Each teacher was given a hand tally counter with which to record the number of responses given during the double lessons. Teachers A and B recorded instances of reprimand during their reprimand phase and instances of praise during the combined phase. The numbers of instances of praise and reprimand recorded during observed lessons were given to the observer on conclusion of the subsequent observed session.

Throughout the study the teachers were asked not to discuss any aspect of the research with colleagues. Eight weeks after the conclusion of the study observers returned to the school and observed teachers A, B and D on three separate occasions with the same class as observed during the main study (class C had left school by this point). During this phase no instructions of any kind were given to the teachers and the self-recording procedures were not implemented.

Inter-observer agreement

OPTIC has repeatedly been shown to be highly reliable, as measured by inter-observer agreement figures, when employed by trained observers. Merrett and Wheldall (1986) reported average inter-observer agreement figures of 93 per cent for teacher behaviour and 94 per cent for pupil on-task behaviour. In the present study a second trained observer was present for one of the observation sessions for each teacher/class during each intervention phase. Inter-observer agreement was calculated in the usual way by dividing number of agreements by the total of agreements and disagreements between the two observers.

RESULTS

All obtained levels of inter-observer agreement for each teacher during each phase (15 in total) of the study exceeded the generally accepted level of 80 per cent except for teacher C during baseline, section A = 75 per cent). Mean levels of inter-observer agreement for section A and section B of OPTIC respectively, were: teacher A 94 per cent and 92 per cent; teacher B 99 per cent and 93 per cent; teacher C 92 per cent and 93 per cent; and teacher D 100 per cent and 95 per cent. This constitutes good evidence for the reliability and objectivity of the obtained data.

Data gathered on rates of teacher approval and reprimand during the baseline and intervention phases are summarised in Table 1. The baseline data indicated clearly that teacher A and B gave more negative than positive comments to their pupils, the vast majority of which were public reprimands to social behaviour. High rates of teacher negative responding and low rates of positive responding were evident in both classrooms. Teachers C and D displayed low rates of both positive and negative responding.

During the private reprimands phase teacher A made dramatic changes in her negative responding. Total reprimanding dropped to only one sixth of baseline levels. This included a reduction in the mean rate of public reprimanding to less than one per observation session. Private reprimanding also decreased to an appropriate level. During the (combined) private reprimand and praise phase teacher A reduced her overall rate of reprimanding still further and doubled her mean rate (overall) for praising. Private praising increased four-fold. During follow up observations teacher A maintained both her low rate of reprimanding and her increased rate of praising. However, her praise to pupils had reverted to being more public in delivery.

Teacher B also made major changes in his behaviour during the private reprimands phase. His overall mean rate of reprimanding decreased to a quarter of baseline levels. Again public reprimanding dropped to less than one per observation session. Conversely, teacher B's mean rate of private reprimanding increased but to an acceptable level. During the combined private reprimands and praise phase teacher B experienced a slight increase in reprimanding. On the other hand, his mean rate (overall) of praise increased slightly. The most marked increase occurred in private praising where the mean rate increased three-fold. During follow-up observations, teacher B reduced his mean rate of reprimanding still further, giving no public reprimands and few instances of private reprimand. However, his rate of praising pupils had declined.

Teacher C more than doubled his rate of praising pupils during the private praise phase. Moreover, his rate of private praising more than doubled. During the combined private praise and reprimands phase teacher C experienced a small reduction in his overall mean rate of praising. This was a result of fewer public approving responses being given. With regard to reprimands teacher C's overall mean rate increased slightly.

Teacher D's overall mean rate of praising increased by over 50 per cent during the private praise phase. The most marked change was evident in private praising where his mean rate increased tenfold. During the combined private praise and reprimanding phase, teacher D experienced a reduction in his overall mean rate of praising, the most noticeable fall being in private praise. On the other hand teacher D increased his overall mean rate of reprimanding, the resultant rate being equally divided between public and private reprimands. During follow up observations teacher D had increased his overall mean rate of praising and decreased the number of reprimands delivered. Most of the remaining reprimands were delivered privately.

The results for class on-task behaviour during baseline and intervention phases are shown graphically for all classes in Figure 1. During baseline observations class A had a mean

TABLE 1

Mean Rates of Teacher Positive and Negative Responding (per 15 Minutes) across Baseline and Intervention Phases.

	POSITIVE			NEGATIVE		
	Public	Private	Total	Public	Private	Total
A	2.0	1.5	3.5	13.5	4.2	17.7
B	4.9	1.8	6.7	9.2	0.5	9.7
C	0.8	2.6	3.4	1.4	0.8	2.2
D	3.8	0.6	4.4	2.2	0.4	2.6

(i) Mean rates of teacher positive and negative responding (per 15 minutes) during baseline observations.

	POSITIVE			NEGATIVE		
	Public	Private	Total	Public	Private	Total
A	0.8	0.4	1.2	0.6	2.2	2.8
B	3.0	1.2	4.2	0.6	1.4	2.0
C	1.4	6.8	8.2	0.2	0.2	0.4
D	0.8	6.0	6.8	0.4	0.4	0.8

(ii) Mean rates of teacher responding during the private reprimands phase (teachers A and B) and private praise phase (teachers C and D). (Target rate of private reprimands fewer than 2.1 for A and 1.7 for B; target rate of private praise at least 4.2 for C and D.)

	POSITIVE			NEGATIVE		
	Public	Private	Total	Public	Private	Total
A	1.2	6.8	8.0	0.8	1.2	2.0
B	2.0	6.6	8.6	0.8	1.6	2.4
C	0.4	6.2	6.6	1.0	2.0	3.0
D	2.4	2.0	4.4	2.2	2.2	4.4

(iii) Mean rates of teacher responding during the private reprimands and private praise (combined) phase for all teachers. Target rate of private praise for A and B at least 3.2; target rate of private reprimands for C and D up to 1.1).

	POSITIVE			NEGATIVE		
	Public	Private	Total	Public	Private	Total
A	5.3	2.3	7.6	1.7	1.7	3.4
B	3.3	1.3	4.6	0.0	0.7	0.7
C		(no data available)				
D	2.7	3.0	5.7	0.6	1.7	2.3

(iv) Mean rates of teacher positive and negative responding during follow up observations.

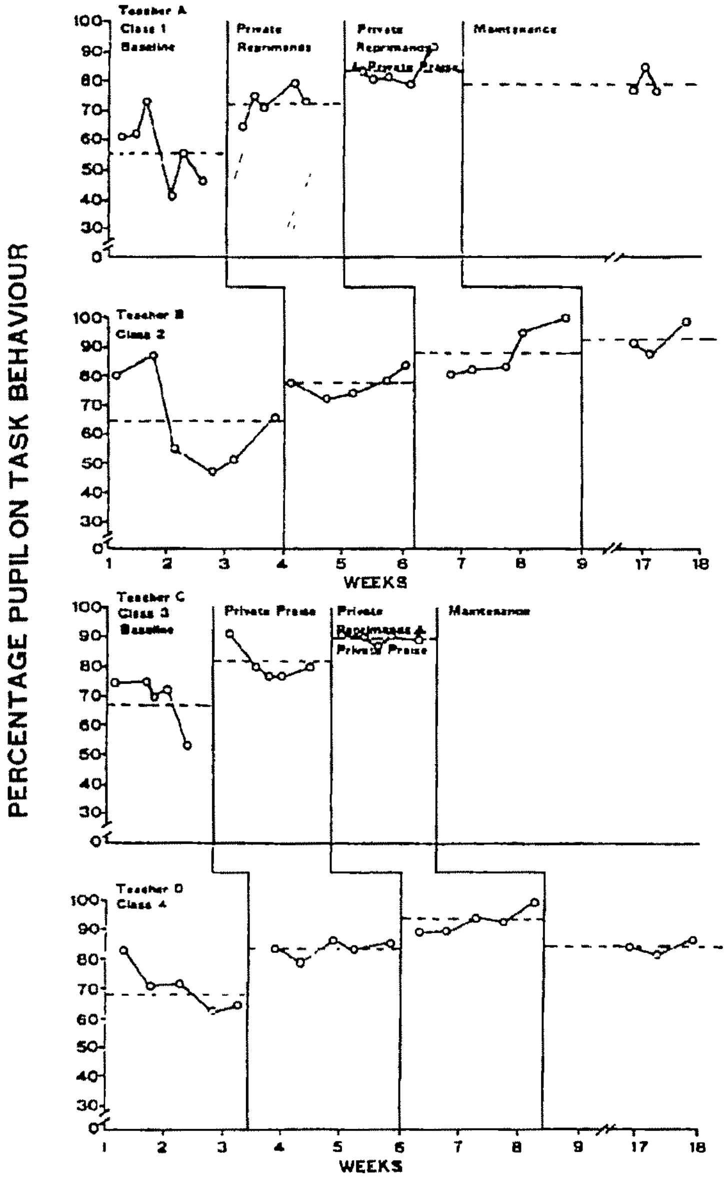
on-task behaviour level of 55 per cent (range 39 per cent to 71 per cent). This increased by 16 per cent to a mean level of 71 per cent (range 62 per cent to 79 per cent) during the private reprimands phase. A further increase of 11 per cent to a mean level of 82 per cent (range 77 per cent to 90 per cent) occurred during the private reprimands and praise phase. During follow up observations the mean on-task behaviour level fell slightly to 77 per cent (range 73 per cent to 82 per cent).

In class B (where identical interventions were introduced one week after class A) mean on-task behaviour averaged 63 per cent (range 47 per cent to 85 per cent) during baseline observations. This increased by 13 per cent to a mean level of 76 per cent (range 72 per cent to 80 per cent) during the private reprimands phase. A further increase of nine per cent to a mean level of 85 per cent (range 79 per cent to 94 per cent) occurred during the private reprimands and praise phase. During follow up observations the mean on-task behaviour level increased slightly to 88 per cent (range 84 per cent to 94 per cent).

During baseline observations class C had a mean on-task behaviour level of 68 per cent (range 52 per cent to 74 per cent). This increased by 13 per cent to a mean level of 81 per cent

FIGURE 1

OBSERVED LEVELS OF CLASS ON-TASK BEHAVIOUR PER SESSION DURING BASELINE AND INTERVENTION PHASES.



(range 78 per cent to 90 per cent) during the private praise phase. A further increase of 7 per cent to a mean level of 88 per cent (range 85 per cent to 89 per cent) occurred during the private praise and reprimands phase. No follow up observations were possible with this class.

In class D (where identical interventions were introduced one week after class C) mean on-task behaviour averaged 69 per cent (range 61 per cent to 81 per cent) during baseline observations. This increased by 13 per cent to a mean level of 82 per cent (range 79 per cent to 84 per cent) during the private praise phase. A further increase of 6 per cent to a mean level of 88 per cent (range 85 per cent to 92 per cent) occurred during the private praise and reprimands phase. During follow up observations the mean on-task behaviour level decreased to 82 per cent (range 79 per cent to 84 per cent).

To summarise, the mean increases in pupil on-task behaviour from baseline to the second intervention phase were: teacher A 27 per cent, teacher B 22 per cent, teacher C 20 per cent and teacher D 19 per cent.

DISCUSSION

In each classroom in this experiment it was demonstrated that change had taken place in the behaviour of both teachers and pupils. This occurred irrespective of the order in which the interventions were applied. Indeed, since private reprimands and private praise were employed separately and then in combination, the results suggest that each variable is independent of the other and that both may be seen as functional.

The data presented show that, broadly, teachers were able to follow the instructions given to them during all phases of the experiment. Teachers A and B almost eliminated their public reprimands and kept under their target of private reprimands. They subsequently easily exceeded their private praise targets. Similarly, C and D easily met their praise targets. They exceeded their private reprimand targets but rates were still low. However, unplanned changes also occurred; for teachers A and B praise also went down when reprimands were reduced and for C and D the already low rates of reprimands declined further during the first intervention. Consequently, we cannot prove unequivocally what caused the increases in class on-task behaviour but the overall pattern of the data strongly supports our interpretation.

The important role played by the hand tally counters was validated by teacher feedback. Teachers reported that the counters were a constant reminder of what to say and how to say it. Furthermore, they were relatively unobtrusive, allowing the teachers to continue with their normal classroom activities. According to Barlow (1980) the involvement of teachers in recording their behaviour is important, since it may prevent a widening of the scientist-practitioner gap and also allows teachers to gain knowledge of how their behaviour influences pupil behaviour.

Each time that private reprimands were introduced pupil on-task behaviour increased. By employing private reprimands teachers were less likely to disturb other pupils who were working, the majority of pupils not even being aware that a peer was, in fact, being reprimanded. This is in line with the finding of O'Leary *et al.* (1970) that soft reprimands did not single out a particular child, thereby making his behaviour noticeable to others. Furthermore, by making fewer public interruptions the teachers allowed pupils to spend more time working on the task in hand. Private reprimands also have the advantage of avoiding public confrontations between teachers and pupils, particularly if the latter is made aware of exactly what is being disapproved of. This accords with the findings of Jones and Miller (1974) who found that by providing a clear explanation of what is being disapproved of, reprimand efficacy is enhanced. It is interesting to note that since the rates at which the teachers in the present study reprimanded their pupils remained low, the long-term effects of the reprimands must also have been increased, again in accord with Jones and Miller (1974).

Private praising of pupils' behaviour appears to have been effective in this experiment, which supports the finding of Houghton *et al.* (1988) that older pupils prefer private praise.

Indeed, knowing that praise will be private (and therefore less embarrassing) pupils are actively encouraged by teacher attention to behave appropriately. It may also be argued that since teachers were looking for behaviour to praise they concentrated less on behaviour to reprimand. Thus, inappropriate behaviour was less likely to be reinforced through teacher attention.

In three of the four classes baseline levels of on-task behaviour were very similar in spite of the variation in rate of teacher praise and reprimand observed. Similarly, in all four classes there is at least some evidence for falling baselines, indicating deteriorating conditions. It appears that minimal teacher feedback is as bad or as ineffective as high frequency loud reprimands in terms of maintaining pupils' on-task behaviour.

It has been argued that reprimands and praise are mutually enhancing (Azrin and Holz, 1966). This would appear to be the case, since the employment of private reprimands and private praise together resulted in further increases in pupil on-task behaviour in all four classrooms in this study (see Figure 1).

Because the teachers in this study recognised improvements in their classrooms as meaningful, it is possible that they too were positively reinforced. In other words, improved pupil behaviour reinforced teacher behaviour. Support for this comes from the follow-up observations. Eight weeks after the programme ceased three teachers observed had to a large extent maintained their behaviour. High levels of pupil on-task behaviour were also maintained.

ACKNOWLEDGEMENTS.—We would like to thank Professor Ted Glynn and Dr Frank Merrett for their helpful comments on an earlier draft of this paper.

Correspondence and requests for reprints should be addressed to Professor Kevin Wheldall, Director, Special Education Centre, School of Education, Macquarie University, Sydney, NSW 2109, Australia.

REFERENCES

AZRIN, N. H., and HOLZ, W. C. (1966). In HONIG, W. K. (Ed.), *Operant Behaviour: Areas of Research and Application.* New York: Appleton.

BARLOW, D. H. (1980). Behaviour therapy: the next decade. *Behav. Therapy*, 11, 315-328.

HOUGHTON, S., MERRETT, F., and WHELDALL, K. (1988). The attitudes of British secondary school pupils to praise and reward, reprimands and punishment. *N.Z J. educ. Stud.*, 23, 203-214.

HOGAN, W. A., and JOHNSON, D. P. (1985). Elimination of response cost in a token economy program and improvement in behaviour of emotionally disturbed youth. *Behav. Therapy*, 16, 87-98.

JONES, F. H., and MILLER, W. H. (1974). The effective use of negative attention for reducing group disruption in special elementary school classrooms. *Psychol. Rec.*, 24, 435-448.

KAZDIN, A. E. (1984) *Behaviour Modification in Applied Settings.* Homewood, Ill.: Dorsey Press.

KAZDIN, A. E., and KLOCK, J. (1973). The effect of non-verbal teacher approval on student attentive behaviour. *J. appl. Behav. Anal.*, 6, 643-654.

MERRETT, F., and WHELDALL, K. (1986). Observing Pupils and Teachers In Classrooms (OPTIC): a behavioural observation schedule for use in schools. *Educ. Psychol.*, 6, 57-70.

MERRETT, F., and WHELDALL, K. (1987). British teachers and the behavioural approach to teaching. In WHELDALL, K. (Ed.), *The Behaviourist in the Classroom.* London: Allen and Unwin.

MERRETT, F., and WHELDALL, K. (1988). *The Behavioural Approach to Teaching Secondary Aged Pupils (BATSAC).* Birmingham: Positive Products.

O'LEARY, D., and O'LEARY, S. G. (1977). *Classroom Management: The Successful Use of Behaviour Modification.* (2nd ed.), New York: Pergamon.

O'LEARY, K. D., KAUFMAN, K. F., KASS, R. E., and DRABMAN, R. S. (1970). The effects of loud and soft reprimands on the behaviour of disruptive students. *Except. Children*, 37, 145-155.

O'LEARY, S. G. (1985). It's time to stop avoiding punishment. Invited address to the Association for the Advancement of Behaviour Therapy Convention, Houston, Texas.

SCOTT, L., MCNAMARA, E., and MCPHERSON, E. (1986). The use of behaviour modification in the secondary school: a further development. *Educ. Child Psychol.* 3, 4-20.

VAN HOUTEN, R., NAU, P. A., MACKENZIE-KEATING, S., SAMEOTO, D., and COLAVECCHIA, B. (1982). An analysis of some variables influencing the effectiveness of reprimands. *J. appl. Behav. Anal.*, 11, 91-94.

WHELDALL, K., and AUSTIN, R. (1980). Successful behaviour modification in the secondary school — a reply to McNamara and Harrop. *British Psychological Society Division of Educational and Child Psychology*

Occasional Papers, 4, (3), 3-9.

Wheldall, K., Houghton, S., and Merrett, F. (1989). Natural rates of teacher approval and disapproval in British secondary school classrooms. *Br. J. educ. Psychol.*, 59, 38-48.

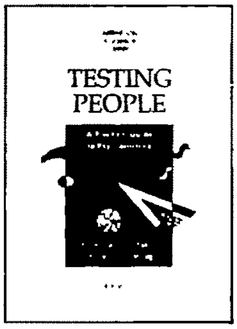
Wheldall, K., and Merrett, F. (1985). Reducing troublesome classroom behaviour in the secondary school. *Maladj. Therap. Educ.*, 3, 37-46.

Wheldall, K., and Merrett, F. (1987). What is the behavioural approach to teaching? In Hastings, N. and Schwieso, J. (Eds.), *New Directions in Educational Psychology II: Behaviour and Motivation in the Classroom*. London: Falmer.

Wheldall, K., Merrett, F., and Borg, M. (1985), The Behavioural Approach to Teaching Package (BATPACK): an experimental evaluation. *Br. J. educ. Psychol.*, 55, 65-75.

(*Manuscript received 24th November*, 1989)

All you need to know about psychometrics



TESTING PEOPLE tackles the fundamental issues and goes on to provide invaluable practical advice on:

- Choosing the best test
- Interpreting data and results
- The thinking behind test construction
- Computer-assisted assessments

£25.00 Hardback 0 7005 1255 1
£15.00 Softback 0 7005 1256 X

TESTING PEOPLE is the first in a brand new series; the NFER-NELSON Assessment Library helps existing test users to increase their expertise and introduces other professionals to the possible uses of assessment.

For further information, or to order, contact our Customer Support Department on Windsor (0753) 858961.

NFER-NELSON, Darville House, 2 Oxford Road East, Windsor, Berkshire SL4 1DF.

Br. J. educ. Psychol., 60, 266-283, 1990

THE DEVELOPMENT OF CONCEPTS ABOUT WRITTEN LANGUAGE IN THE FIRST THREE YEARS OF SCHOOL

By JOAN MASON
(School of Education and Language Studies, Macarthur College, University of Western Sydney, NSW)

SUMMARY. The purpose of this study was to test a model of children's acquisition of concepts about written language in the early school grades. The relationships of a wide range of written language concepts that previous studies suggested are related to children's reading acquisition were examined and developmental trends were assessed in a cross-sectional study of 189 children from kindergarten, grade one and grade two. Factorial analyses were applied to test a three-component model of 14 distinguishable, logical constructs underlying young children's concepts about written language. The three components were concepts about the functions and purposes of written language, concepts about the technical features and conventions of written language, and concepts about the processes and strategies of reading. Results suggest that the three subtests used are valid and reliable measures of the constructs underlying the children's developing understanding about written language as they learn to read.

INTRODUCTION

THE beginnings of written language, like oral language, lie in children's perceiving some of its functions and purposes for themselves and, therefore, seeking to master its conventions. An awareness of what the task involves is an integral part of reading acquisition. This "task awareness" (Downing, 1984) involves the child in gradually becoming conscious of how and why reading and writing occur. The child becomes aware of the goal to be reached, and of the success or otherwise of the actions carried out to attain that goal. A conceptual representation of at least some of the properties of print, some of the actions readers and writers perform, and some of the features of print is then constructed by the learner. With each experience involving written language use, knowledge is modified and other aspects of the task are focused upon.

There have been a number of investigations of the relationship between development of various concepts about written language and reading acquisition (Evanechko *et al.*, 1973; Ayres *et al.*, 1977; Clay, 1979b; Day and Day, 1980; Johns, 1980; Mason, 1980; Lomax and McGee, 1987). These researchers have substantiated the relationship between several dimensions of children's knowledge about written language and reading growth. The dimensions include an awareness of the functions of written language, recognition of the graphic representation of language, awareness of the relationship between spoken and written language and knowledge of grapho-phonemic relationships. However, while a number of measures of children's awareness of written language have been devised, they do not include all the dimensions that have been found to relate to reading development. The intention of this study was to devise a set of measures representing these dimensions that could be used to test the development of concepts about written language of groups of children in the early school grades.

Previous research was extended in two ways. First, the relationships of a wide range of written-language concepts were examined and developmental trends were assessed in a cross-sectional study of children from kindergarten to grade two. Second, factorial analyses were applied to test a three-component model of 14 distinguishable, logical constructs underlying young children's concepts about written language.

Developmental model of concepts about written language

The first component of the model, *concepts of the functions and purposes of written language*, involves the ability to determine how words and ideas are shared by speech and print. It includes five logical constructs: recognition of reading behaviour; recognition of writing behaviour; awareness that written language conveys a message; awareness of purposes for reading; and awareness of purposes for writing. Mason (1984), Downing *et al.* (1978) and Huba and Kontos (1985) among others, have advanced evidence from the findings of their investigations to support the argument that awareness of the communicative function of written language is the basis of learning about the relationships between spoken and written language. While some awareness of the communicative purpose of written language tends to emerge quite early in the pre-school years (Mason, 1980; Heibert, 1981; Lomax and McGee 1987), active participation in various reading and writing situations appears to be necessary in the development of awareness of the function and purposes of written language. It has also been noted that variability in awareness of the functions and purposes of written language may have an effect upon the child's ability to develop other concepts which relate to subsequent reading acquisition (Ayres *et al*, 1977; Ayres and Downing, 1982; Huba and Kontos, 1985). Moreover, results of a longitudinal study of children over a period of three years from the time of entry into school by Clay (1970) suggest that the effect of these early differences may remain even after all children have acquired the concepts.

The second component, *concepts about the technical features and conventions of written language*, describes the social and task constraints of using written language. It involves understanding the langauge used to talk about written language as well as the arbitrary features of written language. Goodman and Goodman (1979) note that as children respond to written language in its various contexts they almost simultaneously begin to respond to significant features, and may even use metalinguistic terminology to suggest their developing rule structures. Therefore, it is assumed that an awareness of some of the functions and purposes of written language would stimulate awareness of some of the concepts in this component which comprises five logical constructs: awareness of the letter forms, recognition of punctuation, awareness of the directionality of print, recognition of a word in text, and awareness of the language used to describe the features of print. A large number of investigations have found that knowledge of the features of written language is related to learning to read (Reid, 1966; Downing, 1970; Francis, 1973, 1982; Ayres *et al.*, 1977; Clay, 1979b; Haussler, 1984; Huba, 1984), and Francis (1973) found a high correlation between the reading achievement of first grade children and knowledge of the terminology used to describe features of print, even with general vocabulary knowledge controlled.

The third component, *concepts about the processes and strategies of reading*, describes an explicit awareness of the use of visual and phonological features of written language. It involves an awareness of how the stream of oral language may be broken into abstract phonemic grapho-phonemic units that correspond to words and letters. Four logical constructs are included in this component: children's awareness of the variation in letter patterns in written words; awareness of grapho-phonemic correspondences; awareness of the match between spoken and written language and word sequence; and orthographic awareness (spelling to sound relationships). Clay (1979a), Francis (1973, 1982) and Mason (1980) among others, found that the ability to track and discriminate words as discrete visual patterns in a line of print necessarily commenced before children were able to begin relating sounds to letters in reading. Ehri and Wilce (1985) noted that the difference between children who were not reading and those who had just begun to read was that the beginning readers were attempting to relate sounds to letter symbols while the non-readers were using visual patterns only. The observations of Mason (1980), Francis (1982) and Morris (1983) all suggest that word reading indicates some knowledge of letter sound associations. Also, a number of researchers (Baron, 1977; Mason, 1977; Venezky and Massaro, 1979) have found that knowledge of graphemic and phonological conventions relate to awareness of orthographic structures which involve visual recognition of regular patterns of letter clusters within words. Orthographic knowledge involves visual recognition of regular patterns of letter clusters within words. Venezky and Massaro (1979), in a review of a number of experiments involving word recognition, concluded that there is a high positive correlation between orthographic knowledge and word recognition.

This study sought to investigate the validity of this set of constructs and to elucidate the factors present within a wide range of concepts about written language and the relationship between them in a cross-sectional study of children from kindergarten to grade two for each of the component measures. It was also planned to report on the reliability of a set of group administered tests containing items designed to evaluate the development of concepts about written language by school beginners in the early phases of literacy development.

METHOD

Sample

The study was conducted in May, four months after the commencement of the school year. One hundred and eighty-nine English-speaking children of middle-lower socio-economic status enrolled in nine classes from four schools in the western and south western regions of the metropolitan area of Sydney comprised the sample. The schools were not randomly selected, but had indicated an interest in developing a reliable and easily administered group test for predicting reading progress in the first years of school. There were 47 kindergarten children (mean age 5 years 5 months), 80 first grade children (mean age 6 years 4 months) and 62 second grade children (mean age 7 years 5 months). Prior to commencing school, none of the kindergarten children had attended pre-school although a few had attended day-care centres in the previous year. All the kindergarten children were described by their teachers as still being in the "pre-reading" phase.

Instrument

A group survey test comprising three subtests was devised for use in this study. Items from the Linguistic Awareness in Reading Readiness (LARR) Tests (Ayers *et al.*, 1977) which had been found to be effective predictors of subsequent reading achievement for grade one Canadian children (Downing *et al.*, 1978; Ayers and Downing, 1982) were used in the first two subtests. The third subtest was similar in presentation to the LARR material and was designed to test more advanced concepts about the processes involved in reading text.

Subtest 1: The functions and purposes of literacy

Twenty-two of the 44 items from the LARR Test 1: Recognising Literacy Behaviour and 20 of the 40 items in Test 2: Understanding Literacy Functions were taken to make up Subtest 1 in this study. Each item is one or a set of four simple black line drawings which represent various everyday objects and activities, many of which involve reading and writing. In response to the tester's instruction, "Put a circle around the picture that. . .", children are required to indicate the appropriate illustration/s or section of an illustration by drawing a circle around it in response to the test question. Because previous studies (Downing *et al.*, 1978; Swanson and Mason, 1984) noted that the large number of items testing each concept in the LARR tests tended to make them unnecessarily long for young children, some redundant items were eliminated. Four items testing each of the concepts were retained and made into one subtest containing 26 items. This subtest became Subtest 1 intended to tap the following concepts:

> recognition of the reading act; recognition of the writing act; recognition of difference between print and illustrations; awareness that written language conveys a message; awareness of various purposes for reading; awareness of various purposes for writing.

Subtest 2: Concepts about the technical features of written language

The LARR Test 3: Concepts about the Technical Features of Written Language became Subtest 2. In this test some minor modifications were made to illustrations and the wording of a number of items where they were judged to be ambiguous or worded inappropriately for Australian subjects. The test consisted of 30 items intended to tap the following concepts:

> concept of a letter (both upper and lower case), word, number, sentence; punctuation (full stop, question mark, inverted commas, comma); knowledge of directionality on a page, in a line of text, or a word (left to right, first, last, top and bottom).

Subtest 3: Concepts about the process of reading

The third subtest was designed to measure concepts about the processes of reading not included in the LARR test. It contained items which measure the ability to track words in a line of print, word discrimination within a text, grapheme-phoneme knowledge, and orthographic awareness. A number of researchers have found that knowledge of these and similar concepts are predictors of subsequent reading achievement (Francis, 1977, 1982; Clay, 1979a, 1979b; Day and Day, 1980; Johns, 1980). Some sample items from this subtest may be found in the appendix to this paper. Subtest 3 consisted of 22 items intended to tap the following concepts:

> awareness of variation in letter patterns (visual word matching, visual matching of letter sequences in words); awareness of the match between spoken and written language (matching spoken to written words, rhyming words, awareness of oral and written words sequence in sentences); grapho-phonic awareness; and orthographic awareness.

Procedure

Testers who were practising infant grade teachers tested the children in grade groups of no more than 12 children with another person present to assist in supervision. Each subtest was given in the morning on each of three days to avoid tiring the children. While all kindergarten children seemed comfortable when following instructions and selecting item answers in Subtests 1 and 2 a considerable number of children became confused about the task when attempting to complete Subtest 3 which tested more complex print concepts. They required a good deal of further explanation, personal attention and reassurance although there had been no evidence of this behaviour when Subtest 1 or Subtest 2 were administered and the format and procedures for Subtest 3 were very similar. It was therefore decided that the performance of kindergarten children should only be assessed on Subtests 1 and 2 which measured knowledge of the functions and purposes of written language, and knowledge of the technical features and language of reading instruction. Grade one and grade two children were able to cope comfortably with all three subtests.

Scoring

Using the scoring procedure of the LARR test recommended by the authors, Ayers *et al.* (1977), children were given a mark only if the whole item was correctly answered. No mark was awarded for partially correct items. For example, if two objects could be circled as part of the correct answer, a child was given no credit for circling only one, even if no incorrect item were circled. All tests were marked twice by different markers to ensure accuracy in scoring and totalling of scores.

RESULTS

The results were first analysed to reveal developmental trends of children's concepts about written language across the grades in the first three years of schooling.

Reliability of tests

The Kuder-Richardson formula 20 was used to determine the internal consistency among test items. Reliability coefficients are shown below in Table 1. Kindergarten reliability coefficients for Subtests 1 and 2 are high and similar to those obtained in the trial of the revised LARR test reported by two of the authors (Ayers and Downing, 1982). Reliability coefficients were calculated on each subtest for each grade after items that had zero variance were omitted. The reliability coefficients for grades one and two in Subtest 3 indicate a high internal consistency among the test items for that subtest for both grades.

TABLE 1

RELIABILITY COEFFICIENTS OF SUBTESTS, 1, 2 AND 3 BY GRADE

		Subtest 1 (26 items) r	Subtest 2 (30 items) r	Subtest 3 (22 items) r
Kindergarten	(N=47)	0.91	0.88	—
Grade One	(N=80)	0.70	0.65	0.91
Grade Two	(N=62)	0.67	0.66	0.81

As anticipated, the reliability coefficients for grades one and two in Subtests 1 and 2 were considerably lower than those of the kindergarten children. Grade one results and grade two results in Subtests 1 and 2 yielded a number of items with zero or very low variances when items were answered correctly by almost all students, and estimates of the internal consistency of the subtests for these grades are assumed to have been depressed by the consequent reduction in test variance and item intercorrelation values. A similar, though less marked effect, occurred in the Subtest 3 results for grade two.

Grade performance

Table 2 presents the grade mean scores and standard deviations on each of the three subtests.

TABLE 2

GRADE MEANS AND STANDARD DEVIATIONS ON SUBTESTS 1, 2 AND 3

		Subtest 1 (total score=26)		Subtest 2 (total score=30)		Subtest 3 (total score=22)	
		Mean	SD	Mean	SD	Mean	SD
Kindergarten	(N=47)	14.55	6.81	15.66	6.39	—	—
Grade One	(N=80)	19.09	3.28	24.34	3.06	14.53	5.71
Grade Two	(N=62)	20.27	2.91	27.42	2.39	20.31	2.39

An analysis of variance, which was employed to determine whether the pupils in the three grades differed significantly in their performance in each of the three subtests, revealed a significant difference between the mean scores of grades for each of the subtests: Subtest 1 ($F = 25.27$, df 2, 186, $P<0.001$), Subtest 2 ($F = 121.75$, df 2, 186, $P<0.001$), and Subtest 3 ($F = 55.92$, df 1, 140, $P<0.001$).

However, there were problems with the interpretation of the ANOVA because the lack of homogeneity of within-cell variances which was not removable by standard forms of data transformation. Therefore, chi-square comparisons of frequency distributions were used (Kirk, 1968) to confirm the relationship between grade and score on concept subtests.

Chi-square analysis of paired grades, shown in summarised form in Table 3, yielded significant differences between adjacent grades in each of the subtests. The contingency coefficients confirm the substantial association between score levels and grades.

TABLE 3

CHI SQUARE ANALYSES OF PAIRED GRADES BY CONCEPTS SUBTEST SCORES

	Kinder vs Grade One				Grade One vs Grade Two			
	Chi^2	df	P	c.c.	Chi^2	df	P	c.c.
Subtest 1	42.34	24	0.012	0.50	26.42	16	0.048	0.40
Subtest 2	73.88	26	0.000	0.61	43.56	14	0.0001	0.48
Subtest 3	—	—	—	—	63.21	18	0.000	0.56

c.c. = contingency coefficient

The nature of the relationship between grade and concept scores is further illustrated in Table 4 which sets out the median score values and shows an increase from kindergarten to grade two for the percentage of children in each grade above the median score for the total sample, confirming the substantial association between score levels and grades.

TABLE 4

RELATIONSHIPS BETWEEN GRADE LEVEL AND MEDIAN SCORE FOR THE TOTAL SAMPLE ON EACH SUBTEST

	Chi^2	df	P	c.c.	Median No. of items	% total sample ≥ median score		
						Kinder (N 47)	Grade 1 (N 80)	Grade 2 (N 62)
Subtest 1	74.44	46	0.005	0.53	19/26	40%	51%	76%
Subtest 2	151.30	54	0.000	0.67	24/30	6%	65%	95%
Subtest 3	63.21	18	0.000	0.56	19/22	—	28%	85%

c.c. = contingency coefficient

The results confirm that there is a developmental pattern in the acquisition of concepts about written language and that the reliability of the tests used to measure their acquisition in this study is acceptable. Of particular interest is the evidence that in the early part of the kindergarten year, when most children were unable to recognise even a few familiar words, they already had some knowledge of the purposes and functions of print, and were aware of some of the conventions of print. These concepts continue to develop during the following years while new concepts emerge. It should also be noted, however, that one third of the 47 kindergarten children knew less than 50 per cent of the concepts related to the purposes and functions of written language measured in Subtest 1. Moreover, even after one year of school attendance, grade one children's responses to items in Subtest 1 indicated that a small number of children (8 per cent) were still uncertain about the difference between the act of reading and the act of writing and of the difference between drawing a picture and writing something. It suggests that a large minority of children may enter kindergarten with a low level of understanding about the functions and purposes of written language which could affect their ability to cope with the instructional tasks related to reading and writing and to interpret some of the instruction register used when written language is the subject of discussion.

Evidence of the difference between the level of children's concepts about written language in the early school grades reflects the findings of Day and Day (1980) and Johns (1980) who tested beginning readers using the Concepts about Print Test (Clay, 1979b) and found significant differences between the children in their knowledge of concepts about the conventions used in written language

Development patterns in written language concepts

The major focus of this part of the study was to examine the usefulness of a three-component model of logically distinguishable functions in detailing children's development of concepts about written language in the first years of schooling.

A factor analysis of the results was carried out in order to examine patterns and inter-relationships in the development of concepts about literacy during the first three years of school instruction. The items included in the three subtests in this study were selected to measure logically distinguishable concepts based upon research discussed earlier, as summarised below:

Subtest 1: Concepts of the functions and purposes of written language

(i) recognition of written language
(ii) recognition of reading behaviours
(iii) recognition of writing behaviours
(iv) awareness of purposes for reading
(v) awareness of purposes for writing

Subtest 2: Concepts about the technical features and conventions of written language

(i) directionality
(ii) discrimination of letters
(iii) recognition of a word
(iv) awareness of punctuation
(v) awareness of metalinguistic terms

Subtest 3: Concepts about the processes of reading

(i) awareness of letter patterns
(ii) awareness of match between spoken and written language
(iii) grapho-phonic awareness
(iv) orthographic knowledge (letter order)

The results of the factor analyses based on kindergarten, grade one and grade two samples are presented below. These analyses examined the responses of the sample of 47 kindergarten, 80 grade one and 62 grade two subjects in order to ascertain whether there were identifiable factors within these subtests and whether they included a range of items that measure the concepts intended.

Separate factor analyses of the item correlation matrices for Concepts Subtests 1 and 2 were undertaken for each of two groups, kindergarten and grade one. Grade two data from Subtests 1 and 2 were not factor analysed as a large number of the items were answered correctly by most grade 2 children. The item correlation matrices of grade one and grade two for Concepts Subtest 3 were subjected to separate factor analyses. Kindergarten children had not been tested with this material.

The maximum likelihood method of factor analyses was used to determine the minimum number of factors needed to explain the correlations among the items. Oblimin rotations were used as it was expected that factors could be correlated. Factors are defined in terms of those items with factor loadings of ≥0.40 for kindergarten (N=47) and ≥ 0.30 for grade one (N=80) and grade two (N=62), as loadings of this magnitude are significantly different (P≤ 0.01) from zero.

Subtest 1: Concepts of the functions and purposes of literacy

A maximum likelihood factor analysis of the responses of 47 kindergarten pupils in Subtest 1, which measured concepts about the functions and purposes of literacy, provided five factors. The responses of 80 grade one students to the items in Subtest 1 also provided five factors but provided no minimum, so the six-factor solution which provided a chi square of 178.32 (P=0.226) was accepted while recognising that five factors might be sufficient. Item 4 was omitted from the grade one analysis as the item was answered correctly by all children.

The rotated factor pattern matrices for the two solutions obtained using the kindergarten and grade one Subtest 1 results are presented in summary in Table 5. Factor loadings of the items are listed in italics for the factors identified in the kindergarten group. To show how these factors split in the grade one solution, factor loadings are presented in plain type. Where items do not have any significant loading, the highest actual loading has been included in brackets.

In kindergarten the items that load on Factor 1 suggest that it describes recognition of environmental print. Factor 2 is loaded with items which tap knowledge that handwritten material as well as printed material may be read by others. This factor can be regarded as an awareness that written language has different forms. The items which load on Factor 3 test knowledge of the purposes for reading and writing such as enjoying a story, looking for a telephone number, advertising a sale, or leaving a message. Items loading on Factor 4 define it as recognition of the act of writing. In these items subjects are required to select illustrations of people who are writing and reject those of people painting or drawing. In Factor 5 the items tap an awareness that the purpose of written language is to communicate a message and that this is different from communicating pictorially.

In the grade one analysis, the components of Factor 1 in kindergarten representing a discrimination of environmental print split into three separate factors. The items loading on Factor 1 in grade one clearly define this factor as an ability to recognise environmental print. Factor 2 is loaded with items which are intended to tap awareness of various situations in which people read and this factor is best described as awareness of reading behaviour. The items with the most significant loadings on Factor 3 indicate that this factor is tapping children's knowledge that there is a difference between the functions of writing and illustrating. Children must indicate that only the print can be read and not the illustration that accompanies it. This suggests that the factor is a recognition of the function of written language. Factor 4 in grade one is loaded with a number of items that also were represented in kindergarten's Factor 2. These items tap knowledge that handwritten material as well as printed material may be read by others so this factor can be described as an awareness that written language has different forms. Factor 5 describes a knowledge of the varied purposes for reading. Factor 6 relates to an awareness that the reason for writing is to record or communicate a message or intention to a reader and is best interpreted as an awareness of the purposes for writing.

TABLE 5

ROTATED FACTOR PATTERN MATRICES FOR KINDERGARTEN AND GRADE ONE FOR SUBTEST ONE: CONCEPTS ABOUT THE FUNCTIONS AND PURPOSES OF WRITTEN LANGUAGE

Item	*Factor 1 Kinder*	Factor 1 Grade 1	Factor 2 Grade 1	Factor 3 Grade 1	*Factor 2 Kinder*	Factor 4 Grade 1	*Factor 3 Kinder*	Factor 5 Grade 1	Factor 6 Grade 1	*Factor 4 Kinder*	*Factor 5 Kinder*
1	*0.63*	0.77		0.30							*0.41*
2	*0.69*	0.72									
3	*0.47*	0.56									
4	*0.77*										
5	*0.54*	0.30				–0.32	*0.44*		0.48		
6	*0.69*										
7	*0.57*			0.70							
8	*0.60*		0.60								
9	*0.63*		0.54								
10	*0.68*		0.75								
11									0.42		*–0.46*
12	*(0.39)*								0.32		
13					*0.77*	0.71					
14					*0.50*	0.92					
15	*(0.39)*			0.77							
16					*0.48*	(0.29)					*–0.42*
17										*0.49*	
18				–0.31						*0.82*	
19							*0.53*	0.39			
20			0.40				*0.42*	0.93			
21							*0.60*	(0.28)			
22							*0.44*		0.47	*–(0.39)*	
23											*0.69*
24							*0.61*		0.40		
25							*0.73*	0.31			
26				0.30					0.54		*–(0.38)*

Note: Factor loadings of the items are listed in italics for the kindergarten group and in plain type for the associated factor or factors into which they split in grade one. Where items do not have any significant loading the highest actual loading has been included in brackets.

Evidently, at the grade one level, children's knowledge of various common functions and purposes of literacy is clearly defined, indicating differences in the level and specificity of concepts about the functions and purposes of literacy between kindergarten and grade one as measured by this subtest.

Subtest 2: Concepts of technical features of written language

Maximum likelihood factor analyses of the responses of the 47 kindergarten children and 80 grade one children in Subtest 2 yielded a four-factor solution for kindergarten, but 10-factor solution for grade one. Factor loadings of the items are listed in italics for the kindergarten group and in plain type for the associated factor or factors into which they split in grade one. A summary of the significant loadings on the factors is shown in Table 6.

In kindergarten, the items loading on Factor 1 define it as the ability to recognise letter or number symbols as distinct from other graphic symbols. Factor 2 taps the child's ability to recognise the difference between a letter and a word. Factor 3 is a punctuation

TABLE 6

ROTATED FACTOR PATTERN MATRICES FOR KINDERGARTEN AND GRADE ONE FOR SUBTEST TWO: CONCEPTS ABOUT THE FEATURES AND TECHNICAL ASPECTS OF WRITTEN LANGUAGE

Item	*Factor 1 Kinder*	Factor 1 Grade 1	Factor 2 Grade 1	Factor 3 Grade 1	Factor 4 Grade 1	*Factor 4 Kinder*	Factor 5 Grade 1	*Factor 3 Kinder*	Factor 6 Grade 1	*Factor 4 Kinder*	Factor 7 Grade 1	Factor 8 Grade 1	Factor 9 Grade 1	Factor 10 Grade 1
1										*0.69*				
2	*(0.39)*								0.44					
3	*0.86*	0.39			0.86									
4	*0.99*	0.30					0.40							
5	*0.55*		0.38											
6	*0.78*	0.83	0.45											
7	*0.68*	1.00												
8	*0.56*		0.41											
9										*0.56*			0.37	
10	*0.51*		0.78											
11														
12			0.39			*–0.43*				*0.45*	0.52			
13											0.98			
14		0.30		0.98										
15	*0.56*			0.98										
16	*0.54*			0.83										
17	*0.55*			0.52										
18					0.30			*(0.39)*						
19				0.48				*–(0.37)*						
20												0.81		
21												0.51		
22						*–0.40*						0.56		
23								*0.82*	0.60					
24								*0.92*	0.90					
25			(0.26)			*–0.92*	–0.86			*–0.43*				
26						*–0.74*	–0.77							
27										*0.40*			0.59	
28													0.80	
29														0.55
30														–0.99

Note: Factor loadings of the items are listed in italics for the kindergarten group and in plain type for the associated factor or factors into which they split in grade one. Where items do not have any significant loading the highest actual loading has been included in brackets.

recognition factor and Factor 4 describes recognition of terms used in written language such as number, word, line, sentence and letter.

By comparison, the factor analysis of grade one Subtest 2 results indicates greater specificity in knowledge of technical concepts in children's second year of schooling. Kindergarten's Factor 1 is split in the grade one analysis and the items load on four separate factors. Factor 1 in grade one is recognition of the difference between number and letter symbols because it describes knowledge of what is meant by the technical terms "number" and "letter" as well as the ability to discriminate between letters and numbers. Items in Factor 2 probe the extent and level of the subject's familiarity with various representations of written language, and so it can be described as recognition of the features of written language. Factor 3 is a word discrimination factor requiring not only an awareness of words as a set of graphemes bounded by spaces, but also knowledge that punctuation symbols always have a separate function although they may appear to be part of the final word or initial word in a sentence. Factor 4 represents grade one children's ability to distinguish between the different functions of written symbols which are graphically similar, for example, an awareness of the difference between 2 and Z and between the function of a full stop and the dot on an "i".

Factor 5 in the grade one analysis corresponds to Factor 2 in kindergarten, indicating knowledge of the difference between letters and words. In Factor 6, the two items relating to the question mark load on a separate factor as they did in the kindergarten Factor 3 as a punctuation recognition factor. Factor 7 is defined by a doublet of items representing the subject's concept of directionality of a page of text. While this factor did not appear in the kindergarten analysis, it cannot be assumed that this concept has not been acquired by subjects in that grade. A likely explanation is that for kindergarten subjects knowledge of directionality in print may have been confounded by the need to know also the meaning of the language terms in the question.

Factor 8 and Factor 9 indicate a clear case of the development of concepts in the 12 months between kindergarten and grade one. The three items in Factor 8 relate to knowledge of upper and lower case print and of what constitutes capital letters when presented either singly or in a sentence. The items loading on Factor 9 indicate that this factor relates to the subject's recognition of a written sentence.

Subtest 3: Concepts about the reading process

In Subtest 3 separate factor analyses of the responses of the 80 grade one children and the 62 grade two children were undertaken. Two items were omitted from the grade one analysis and four items were omitted from the grade two analysis as they were answered correctly by all children. Analyses yielded five factor solutions for both grade one and grade two although only four factors could be clearly defined for either grade. The rotated factor pattern matrices for the two solutions obtained using the grade one and grade two results are presented in summary in Table 7.

In grade one Factor 1 is clearly defined by most of the items testing subjects' ability to visually match words or letters and the factor which may be described as word/letter cluster matching. The items that load on Factor 2 indicate that this factor describes orthographic knowledge. These items all tap various aspects of the knowledge of the relationship between sound and letter patterns within well known words. Factor 3 is defined by four items which require subjects to exercise any grapho-phonic knowledge at their disposal together with syntactic knowledge to match a spoken sentence with the one that has the appropriate word order. In the items that load on Factor 4, there are three sentences that are semantically and syntactically correct and are similar in meaning. The subjects must identify the one spoken by the examiner. This factor taps subjects' ability to use visual-auditory memory to search systematically for and match a specific spoken message with its exact written equivalent. It could be best described as the operation of semantic and syntactic knowledge.

Factor 1 in grade two is very similar to Factor 1 in grade one, testing children's ability to visually match word/letter clusters. Factor 2, which is defined by seven items, requires children to use their knowledge of beginning or ending consonants and therefore could be labelled as grapho-phonic knowledge. The fact that a number of the items from Factor 1 also load on this factor led to the assumption that the grade two children were identifying letter clusters phonetically as well as visually. Whereas in grade one items measuring orthographic knowledge cluster together with items involving the implementation of semantic and syntactic expectations on Factor 2, in grade two items measuring orthographic knowledge cluster with the visual pattern matching items on Factor 2 and Factor 3.

TABLE 7

ROTATED FACTOR PATTERN MATRICES FOR GRADE ONE AND GRADE TWO FOR SUBTEST THREE: CONCEPTS ABOUT THE PROCESSES OF READING

Item	*Factor 1 Grade 1*	*Factor 2 Grade 1*	*Factor 3 Grade 1*	*Factor 4 Grade 1*	*Factor 5 Grade 1*	Factor 1 Grade 2	Factor 2 Grade 2	Factor 3 Grade 2	Factor 4 Grade 2	Factor 5 Grade 2
1					*(0.28)*		0.79			
2	*0.64*					0.87	0.39	0.31		
3	*0.41*					0.31	0.40	0.32		
4	*0.40*					0.42		0.30		
5	*0.43*		*0.35*				0.35	0.34	0.50	
6			*0.31*		*0.74*	0.83				
9				*0.48*			0.70			
10				*0.83*			0.69		0.42	
11	*0.50*			*0.38*				0.75		
12	*0.69*			*0.54*			0.58			0.34
13				*0.34*				0.37		0.33
14				*0.60*				0.66		
15		*0.69*	*0.30*							
16		*0.80*							1.04	
17			*(0.29)*	*0.41*						0.33
18			*(0.29)*	*0.46*					0.74	
19				*0.41*					0.41	
20			*(0.29)*	*0.33*				0.54	0.34	
21				*0.60*					0.45	
22		*(0.29)*	*0.91*							

Factor 3 relates more clearly to the ability to discern appropriate orthographic patterns within words and is best described as orthographic knowledge. The items which best define this factor (11, 13, 14, and 20) provide no clues through the beginning and ending consonants and require analysis of the internal structure of the word before a decision can be made.

Seven items load on Factor 4 in grade two, which is defined by items that involve the knowledge of the link between oral and written language. Two highly significant items tap the children's ability to use their semantic and syntactic expectations as well as word analysis. This factor appears to combine knowledge tapped by Factors 2 and 4 in grade one which describe knowledge of sound-symbol relationships and an awareness of the match between oral and written language units. While the factor is best described as a semantic and syntactic knowledge factor, some of the items loading on this factor suggest that at this stage it also involves knowledge of the confirmatory role of grapheme-phoneme relationships in word identification.

The outcome for each of the logically distinguishable functions is summarised in Figure 1.

FIGURE 1

SUBTEST 1 KNOWLEDGE OF THE FUNCTIONS AND PURPOSES OF LITERACY

Logical Function	Identified Factor	
	Kindergarten	Grade One
i. Recognition of written language	Recognition of environmental print	Recognition of environmental print Awareness of range of environmental print
ii. Recognition of reading behaviours	Awareness of the relationship between reading and writing	Awareness of the relationship between reading and writing
iii. Recognition of writing behaviours	Recognition of the writing act	Awareness of how written lanuage is produced
iv. Awareness of purposes for reading	Awareness of reading and writing	Awareness of purposes for reading
v. Awareness of purposes for writing		Awareness of purposes for writing

SUBTEST 2 TECHNICAL FEATURES AND CONVENTIONS OF WRITTEN LANGUAGE

Logical Function	Identified Factor	
	Kindergarten	Grade One
i. Awareness of letter forms	Recognition of numbers and letters	Differentiation between numbers and letters Discrimination of forms of written language Awareness of different functions of letter, number, punctuation symbols
ii. Directionality	not identified	Directionality
iii. Recognition of words	Recognition of term for word	Recognition of words
iv. Awareness of punctuation	Punctuation recognition	Punctuation recognition Sentence discrimination
v. Awareness of metalinguistic terms	Recognition of difference between letters and words	Recognition of difference between letters and words Recognition of upper and lower case letters

SUBTEST 3 PROCESSING WRITTEN LANGUAGE

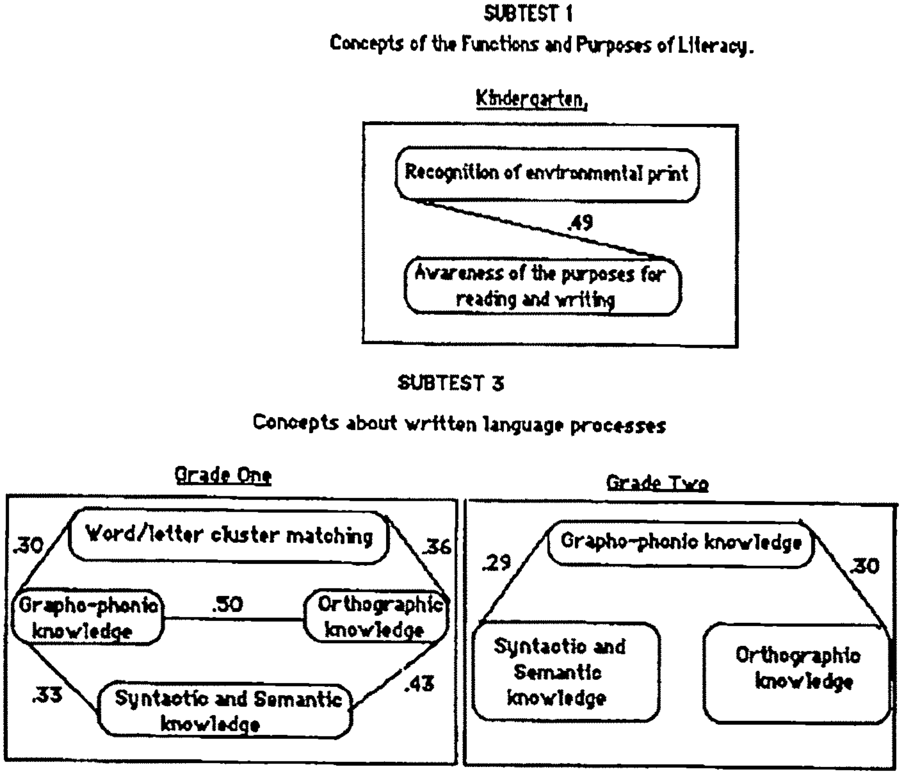
Logical Function	Identified Factor	
	Grade One	Grade Two
i. Awareness of variation in letter patterns	Word/letter cluster matching	Word/letter cluster matching
ii. Grapho-phonic knowledge	Grapho-phonic knowledge	Grapho-phonic knowledge
iii. Awareness of match between spoken and written language	Semantic and syntactic knowledge	Semantic and syntactic knowledge
iv. Orthographic knowledge	Orthographic knowledge	Orthographic knowledge

When examining the correlations between factors, the values used previously (kindergarten≥0.4 and grades one and two≥0.3) were taken as rough indicators that the correlation between factors was significantly different from zero. In the Subtest 1 analysis of the kindergarten results the recognition of environmental print factor was significantly related to awareness of purposes for reading and writing factor (r=0.49). The ability to recognise environmental print is acquired early and the fact that the environmental print concept relates significantly to awareness of the purposes for reading and writing indicates that it may facilitate the development of other concepts through stimulating the child to ask questions and find out more about written language. Gibson (Gibson and Yonas, 1968) has pointed out that the children's ability to discriminate between writing and other graphic forms must be basic to the notion of reading and it is quite likely that this is what the relationship between environmental print concepts and awareness of the purposes of written language represents. After a year of instruction, however, in grade one neither this nor any other factor relates significantly to other concepts about the functions and purposes of written language.

In both the kindergarten and the grade one analysis of Subtest 2, measuring technical features of written language, no significant correlation was found between any of the factors. Four factors in the grade one analysis of Subtest 3, testing concepts of written language processes, were found to be significantly related. Written word/letter cluster matching was correlated to the extent of 0.30 with grapho-phonic knowledge and 0.36 with orthographic knowledge. Grapho-phonic knowledge was also correlated to the extent of 0.50 with orthographic knowledge and 0.43 with orthographic knowledge. A relationship between these factors would be anticipated. Once word boundary awareness is established, the beginning reader is free to develop and utilise phoneme-grapheme knowledge and orthographic knowledge, to use knowledge of syntactic forms to synthesise individual words into meaningful language (Clay, 1979a; Ehri, 1979; Henderson, 1980; Morris, 1983). The relationships are depicted in Figure 2.

FIGURE 2

A Summary of Significant Correlations Between Factors where they Occurred in Each of the Three Subtests in Each Grade.


SUBTEST 1
Concepts of the Functions and Purposes of Literacy.
Kindergarten,
Recognition of environmental print
.49
Awareness of the purposes for reading and writing
SUBTEST 3
Concepts about written language processes
Grade One
Word/letter cluster matching
.30
.36
Grapho-phonic knowledge
.50
Orthographic knowledge
.33
Syntactic and Semantic knowledge
.43
Grade Two
Grapho-phonic knowledge
.29
.30
Syntactic and Semantic knowledge
Orthographic knowledge

This appears to indicate that as the perceptual processing of word-units becomes more co-ordinated and automated the word becomes the frame within which phonemic segments and succession of letter arrangements can be examined. With the stable concept of a word in text established and no longer needing conscious monitoring, it is conceivable that the array of letters within the frame of the word and their correspondence to phonetic patterns become the subject of freed attentional resources.

DISCUSSION

Two results from the analyses require particular attention. Firstly, all logical functions were found within component factors for each grade, excepting directionality in kindergarten in Subtest 2, indicating that these could be regarded as valid indicators of the underlying constructs. Secondly, an examination of the correlations between factors for the different solutions indicates that with few exceptions most factors are largely independent in each grade in Subtests 1 and 2 (Figure 2). As anticipated, the component factors in Subtest 3 tended to be significantly related to each other in grade one although this relationship had lessened in grade two.

Although the sample sizes used in the factor analyses in this study were too small to give a conclusive picture of the concepts that young children develop while learning to read and write in the first school years, the data make it clear that knowledge of all concepts increases significantly over the period of the first two school years.

Two results were revealed in the factorial analyses of the three-component model. First, items measuring the logical functions within each component were found to be valid indicators of the underlying concepts. Second, the factorial analyses demonstrated that even in the first months of school, the kindergarten children in this study displayed some knowledge of each of the logical functions in the two components of the test on which they were assessed. However, their knowledge of these functions appears to be predominantly at the level of recognition of overt behaviours and identification of gross similarities and differences in graphic displays. In the first component, *concepts about the functions and purposes of written language*, they identified written language in the environment, were aware that what was written was intended to be read, could recognise the act of writing, and were able to distinguish reading and writing from other activities. In the second component, *concepts about the features and technical terms of written language*, they could differentiate between numbers and letters and other graphic symbols, and could recognise some of the terms used in talking about written language. The grade one children appeared to have developed a keener awareness of the relationship between behaviour and purpose and displayed increased and more specific knowledge of concepts in the two components. These findings corroborate those of Lomax and McGee (1987) whose structural analyses of a developmental model of reading suggest that recognition of the graphic features of written language and some knowledge of the purposes and functions of written language are precursors of other print concepts relating to the processes involved in reading text.

In the third component, *concepts about the processes of reading*, grade one children displayed an awareness of each of four logical functions. They were able to visually analyse and match word and letter clusters, were aware of the relationship between language structures and written language, and demonstrated some orthographic knowledge and grapho-phonic awareness. The grade two children also showed gains in each of the three components, with a large percentage gaining almost perfect scores in the first two subtests that relate to knowledge of literacy behaviours and the features and technical terminology of written language.

The results of the analyses demonstrate that concepts about the functions and purposes and features of written language begin to be acquired prior to knowledge of how to process print, but that these early-developing constructs are still incomplete when other concepts about the processes of reading also begin to develop. These findings support a developmental model which implies that children will continue to increase in their knowledge of concepts within the hypothesised components and that new concepts may be acquired before earlier concepts are fully mastered (Day and Day, 1979; Lomax and McGee, 1987).

The act of learning is determined by learners' ability to interpret and respond to what they perceive to be the task and the goal to be achieved. The strategies that children develop in order to read are founded upon the understandings about reading that they have and the expectations about how it works which they develop through various experiences with written language. It follows, then, that children interpret and respond to written language tasks and contexts in various ways depending upon their own concepts about written language and its purposes and in the context of the instructional task being presented by the teacher. However, children enter school from various home backgrounds and with very different experiences and notions of literacy. In any classroom, the focus of instruction in various literacy activities is based on an implicit assumption about the understandings a school entrant already has or needs to develop in order to learn to read and write. Therefore, low levels of knowledge about the functions and purposes of literacy upon school entry may affect children's subsequent ability to read if they are placed in instructional programmes based on the assumption that they already have acquired this knowledge.

In a report of longitudinal study which, among other things, examined the relationship between reading achievement and home environment, Wells (1986) found that children who acquired high literacy skills in the first years of school were those who had regularly had stories told and read to them at home. Home-school programmes that encourage parents and teachers to provide frequent opportunities for children to read along with them, and share in stories, songs and rhymes, will expand children's knowledge of books, the language of books, and the conventions of print. Parents and teachers can aim to build children's awareness of the purposes and functions of written language in its various forms, through sharing the many occasions when the need to use written language in context arises — signs, notices, labels, message and letter writing, story book reading and reading to acquire information. Written materials should be produced in contexts that will signal their purposes, encourage their perusal and stimulate children's desire to use them. In this way, children can experience the reality of what is involved in being a reader and writer.

ACKNOWLEDGEMENT.—This study was carried out as part of a doctoral thesis at the University of Sydney. The author wishes to express thanks to Emeritus Professor D. Spearritt of the University of Sydney for his support and assistance, especially in relation to the statistical analyses, and to Dr Lloyd Ollila and Dr Douglas Ayers of the University of Victoria in Vancouver, Canada, for their helpful comments.

Correspondence and requests for reprints should be addressed to Dr Joan Mason, School of Education and Language Studies, Macarthur College, University of Western Sydney, PO Box 555, Campbelltown, NSW 2560, Australia.

REFERENCES

AYRES, D., and DOWNING, J. (1982). Testing children's concepts of reading. *Educ. Res.*, 24, 277-283.

AYRES, D., DOWNING, J., and SCHAEFER, B. (1977). The Linguistic Awareness in Reading Readiness (LARR) Test-Preliminary Edition. University of Victoria, Canada. Revised Edition, Nelson-NFER, 1982.

BARON, J. (1977). Mechanism for pronouncing printed words: use and acquisition. In LA BERGE, D. and SAMUELS, S. J. (Eds.), *Basic Processes in Reading: Perception and Comprehension*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

CLAY, M. M. (1970), *Language Skills: A Comparison of Maori, Samoan and Pakeha Children Aged 5 to 7 years*. Report to the Maori Education Foundation — Part 1, Auckland, May.

CLAY, M. M. (1979a). *Reading: The Patterning of Complex Behaviour*. (2nd ed.) Auckland: Heinemann.

Clay, M. M. (1979b). *The Early Detection of Reading Difficulties.* (2nd ed.) Auckland: Heinemann.

Day, K. C., and Day, H. D. (1979). Development of Kindergarten children's understanding of concepts about printed and oral language. In Kamil, M. L., and Moe, A. H. (Eds.), *Reading Research: Studies and Applications.* Rochester, N.Y.: National Reading Conference.

Day, K. C., and Day, H. D. (1980). *Observations of Kindergarten and First Grade Children's Development of Oral Language, Concepts about Print, and Reading Readiness.* (Publication date 1978.) ERIC document (ED 176 212) January, 1980.

Downing, J. (1970). Children's concepts of language in learning to read. *Educ. Res.*, 12, 106-112.

Downing, J. (1984). Task awareness in the development of reading skill. In Downing, J. and Valtin, R. (Eds.) *Language Awareness and Learning to Read.* New York: Springer Verlag.

Downing, J., Ayers, D., and Schaefer, B. (1978). Conceptual and perceptual factors in learning to read. *Educ. Res.*, 21, 11-17.

Ehri, L. C. (1979). Linguistic insight: threshold of reading acquisition. In Waller, T. G., and MacKinnon, G. E. (Eds.), *Reading Research: Advances in Theory and Practice. Vol. 1.* New York: Academic Press.

Ehri, L. C., and Wilce, L. S. (1985). Movement into reading: is the first stage of printed word learning visual or phonetic? *Reading Res. Q.*, XX, 163-179.

Evanechko, P., Ollila, L., Downing, J., and Braun, C. (1973). Investigation of the reading readiness domain. *Research in the Teaching of English*, 7, 61-78.

Francis, H. (1973), Children's experience of reading and notions of units in language. *Br. J. educ. Psychol.*, 43, 17-23.

Francis, H. (1977), Reading abilities and disabilities: children's strategies in learning to read. *Br. J. educ. Psychol.*, 47, 117-125.

Francis, H. (1982) *Learning to Read.* London: George Allen and Unwin.

Gibson, E. J., and Yonas, P. A. (Eds.) (1968). *The Analysis of Reading Skill.* Final Report, Project No. 5-1213, Cornell University/US Office of Education.

Goodman, K. S., and Goodman, Y. M. (1979), Learning to read is natural. In Resnick, L. B. and Weaver, P. A. (Eds.), *Theory and Practice of Early Reading, Vol. 1.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Haussler, M. M. (1984), *Transitions into Literacy: A Working Paper.* Arizona Center for Research and Development, College of Education, University of Arizona.

Henderson, E. H. (1980). Developmental concepts of word. In Henderson, E. H., and Beers, J. W. (Eds.) *Developmental and Cognitive Aspects of Learning to Spell: A Reflection of Word Knowledge.* Newark, Del.: International Reading Association.

Hiebert, M. M. (1981), Developmental patterns and interrelationships of preschool children's print awareness. *Reading Res. Q.*, XVI, 236-260.

Huba, M. E. (1984), The relationship between linguistic awareness in prereaders and two types of experimental instruction. *Reading World*, 23 (4), 347-363.

Huba, M. E., and Kontos, S. (1985), Measuring print awareness in young children. *J educ. Res.* 78 (5), 272-279.

Johns, J. (1980), First graders' concepts about print. *Reading Res. Q.*, XV (4), 529-549.

Kirk, R. E. (1968). *Experimental Design. Procedures for the Behavioral Sciences.* Belmont, Calif.: Wadsworth.

Lomax, R. G., and McGee, L. M. (1987), Young children's concepts about print and reading: toward a model of word reading acquisition. *Reading Res. Q.*, XXII (2), 237-256.

Mason, J. (1977), *Reading Readiness: A Definition and Skills Hierarchy from Pre-schoolers' Developing Conceptions of Print.* Technical Report No. 59, University of Illinois. Cambridge, Mass.: Bolt, Beranek and Newman.

Mason, J. (1980), When do children begin to read: an exploration of four-year-old children's letter and word reading competencies. *Reading Res. Q.*, XV, 203-227.

Mason, J. (1984), Early reading from a developmental perspective. In Pearson, P. D., Barr, R., Kamil, M. L. and Mosenthal, P. (Eds.), *Handbook of Reading Research*, New York: Longman.

Morris, D. (1983), Concept of word and phoneme awareness in the beginning reader. *Research in the Teaching of English*, 17 (4), 359-373.

Reid, J. F. (1966), Learning to think about reading. *Educ. Res.*, 9, 56-62.

Swanson, B., and Mason, G. (1984). Measuring early reading progress: some non-traditional measures. *J. educ. Res.*, 78, 90-96.

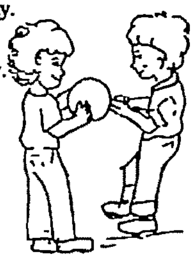
Venezky, R. L., and Massaro, D. W. (1979), The role of orthographic regularity in word recognition. In Resnick, L. and Weaver, P. (Eds.), *Theory and Practice in Early Reading. Vol. 1.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Wells, G. (1986), *The Meaning Makers: Children Learning Language and Use Language to Learn.* Portsmouth, N.H.: Heinemann.

(*Manuscript received 9th November*, 1989)

APPENDIX

SAMPLE OF ITEMS FROM SUBTEST 3

● look cook loud like Jack	4. Put your finger on the little black circle. Put a circle around the word in the row that has an ending just like the one in the box.
✱ her car today fun	12. Put your finder on the black star. Put a circle around the word that sounds almost the same as "gun" . . . it sounds almost the same as "gun".
Boys and grils come uot to paly. Boys nad girls coem out ot play. Boy and girls come out to play. 	13. Put your finger on the picture of the boy and girl playing. Put a circle around the story that says, "Boys and girls come out to play".
I can ride a bike. I can read a book. I can eat a cake. 	18. Put your finger on the girl. Put a circle around the story that goes with the picture.
Jack and hill the Jill up went. Jill the up and Jack went hill. Jack and Jill went up the hill. 	24. Put your finger on Jack and Jill. Put a circle around the story that says "Jack and Jill went up the hill". . . . "Jack and Jill went up the hill".

Br. J. educ. Psychol., 60, 284-298, 1990

FACTORS RELATED TO THE ACADEMIC ATTAINMENTS OF CHILDREN WITH DOWN'S SYNDROME

BY PATRICIA SLOPER, CLIFF CUNNINGHAM, STEPHEN TURNER AND
CHRISTINA KNUSSEN
(Hester Adrian Research Centre, University of Manchester)

SUMMARY. The attainments in reading, number and writing skills, of 117 children with Down's syndrome, aged 6 to 14 years, were assessed using checklists completed by teachers. In a study of child and family functioning a wide range of variables was measured and the relationships of these to academic attainments were investigated using multivariable analysis. The children's mental age scores were most strongly related to academic attainments scores, but, in addition, type of school attended, gender, chronological age and fathers' scores on a measure of locus of control were significantly related. The results are discussed in terms of their implications for educational placement and curriculum, and the role of fathers in their children's education.

INTRODUCTION

IN recent years considerable interest has been shown in the attainments of children with Down's syndrome in the traditional academic skills, particularly in reading (e.g., Buckley, 1985; Lorenz *et al.*, 1985; Casey *et al.*, 1988). Despite this, representative data on levels of attainment are sparse. Some single case studies are reported (e.g., Duffen, 1979), and studies using group data have often been confined to small, selected groups who were able to achieve above the basal level on standard reading or number tests (e.g. Dunsdon *et al.*, 1960; Pieterse and Treloar, 1981). Three studies provide data on more representative samples. Carr (1988) reports data on 41 people with Down's syndrome for whom reading and number tests were administered at the age of 21 years, as part of a longitudinal study. She notes that only two-fifths of these were able to gain reading ages on a standard reading test. "Reading levels" were computed for those who were able to name at least some letters but did not score on the reading test. Three-quarters gained a score on the arithmetic test. It appears, therefore, that wider measures than standard tests, including pre-reading and number levels, are needed to investigate the full range of attainment in children with Down's syndrome.

Buckley and Sacks (1987), in a study of 90 children with Down's syndrome aged 11 to 17 years, report data on reading, writing and number skills, in terms of percentages of children attaining different levels of skills, e.g., can read sentences, can write own name, can do simple addition. Overall scores were computed for school achievement and the relationship between those scores and family position, social class and school was investigated. Significant effects of social class were found, showing advantages for children from non-manual backgrounds, on reading, number and overall scores.

Shepperdson (1988) reports data on the reading and number skills of 51 teenagers, noting that, although there was a strong relationship between reading ability and scores on the Reynell Language Scale, this relationship was not wholly consistent, as some teenagers with high scores on the Reynell scale could not read at all. She suggests that teenagers attending schools for children with moderate learning difficulties were more likely to have mastered academic skills, and that there were some children for whom it seemed likely that they could only read a little because they had not been taught. However, no data are provided to support these suggestions.

Thus, existing studies point to considerable variability in attainment and it appears that this may not wholly be accounted for by differences in mental age or IQ. Whilst IQ may be a strong predictor of academic attainment, as in the general population (e.g., Pellegrini *et al.*, 1987), further investigation of which other factors relate to attainment can provide important information for intervention and education programmes. Evidence on the links between IQ and family factors, such as parental education and social class, for children with Down's syndrome is conflicting (Carr, 1985). It may be that such factors show stronger links with attainments.

Studies which have investigated the factors related to the academic attainments of people with Down's syndrome after controlling for IQ are few. There is some evidence from such studies of higher levels of attainment for children in mainstream schools compared with those in special schools in number work (Casey *et al.*, 1988) and reading (Lorenz *et al.*, 1985); females compared with males, those from non-manual occupational backgrounds compared with those from manual backgrounds, and home-reared compared with non-home-reared in both reading and number skills (Carr, 1988).

In the non-handicapped population links have been shown between social and family factors and reading and number attainment (e.g., Rutter *et al.*, 1970; Hewison and Tizard, 1980; Marjoribanks, 1987). A number of studies have found home environment variables such as cohesion, conflict and engagement in cultural and recreational activities to be related to school achievement in both handicapped and non-handicapped samples (Nihira *et al.*, 1984). Evidence from studies of children with disabilities increasingly points to the importance of family attitudes and involvement in relation to the children's school achievements (e.g., Nihira *et al.*, 1983; Bodner-Johnson, 1986).

However, there may be considerable intercorrelation between such factors as type of schooling and social class, and the delineation of the most important factors requires studies of representative samples, measuring social, demographic and family factors and using multivariable statistical methods. The present study aimed to measure the attainments of a representative sample of children with Down's syndrome, aged 7 to 14 years, in reading, number and writing skills, and to ascertain the factors related to levels of attainment, over and above the effect of mental age. Multiple regression analysis was used to investigate the relationship between attainment and a wide range of social, demographic, child and family variables and to delineate the independent effects of such variables.

METHOD

Sample

The sample consisted of families in the Manchester Down's Syndrome Cohort. This cohort of 181 families of children with Down's syndrome, born between 1973 and 1980, represents 90 per cent of all such births in the Greater Manchester area over this period. In the early years the families participated in a home-based early intervention programme (Cunningham, 1986). At the time of the present study 127 families remained in the cohort. Attenvation had been entirely due to deaths of children and families moving away from the area. Four families declined to participate in the present study, giving a total of 123 families in the sample.

Data were collected through standardised assessments of the children carried out by research staff, checklists completed by teachers, and self-report questionnaires and interviews with mothers and fathers. Child assessment data and teacher checklists were obtained on all children, 121 mothers (99 per cent) were interviewed and 118 (97 per cent) returned questionnaires, 97 fathers (86 per cent) were interviewed and 87 (77 per cent) returned questionnaires. The data reported in this paper are based on the 118 children whose mothers returned questionnaires. The social class distribution of this sample based on current or last occupation of the head of the household compared with national statistics (OPCS, 1987) is given in Table 1. The sample shows a slight over-representation of social classes I, II and III non-manual compared to national figures, but this is not significant (χ^2=3.75, df=3, NS).

TABLE 1

Social Class Distribution (Percentages)

	I	II	III non-man	III man	IV	V
Down's syndrome cohort*	8 (9)	26 (31)	19 (22)	32 (38)	14 (16)	2 (2)
1985 National figures**	7	21	17	38	13	4

Numbers in parenthesis

* Based on Registrar General Classification of Occupations (1980) for father's present occupation where currently unemployed, and mother's current or last occupation if single parent mother.

** For economically active males aged 25-64.

Eleven families (9 per cent) were headed by a single parent. Ten families (8 per cent) were foster or adoptive families, the remainder consisted of at least one natural parent. In 20 families (18 per cent) the father was unemployed at the time of the study. Fifty-eight mothers (49 per cent) were employed, 43 (36 per cent) working part-time.

The mean maternal age at the time of the study was 39.1 years, range 23 to 55; mean paternal age 40.9 years, range 27 to 62; mean child age 9.2 years, range 6 to 14. Sixty-three per cent of the children were male. The mental age of the children was assessed by members of the research team at the time of the study, using the McCarthy Scales of Children's Abilities (McCarthy, 1972) and, for six children who were not able to obtain a basal level on this test, the Bayley Scale of Infant Development (Bayley, 1969). Mean mental age was 43.3 months, range 8 to 90 months, SD 12.7.

The children attended a total of 60 different schools. Thirteen children (11 per cent) attended ordinary classes in mainstream schools; eight (7 per cent), units in mainstream schools; 18 (15 per cent), MLD schools; 76 (64 per cent), SLD schools; one child, an assessment unit; one, a school for the deaf and one was educated at home, by his parents.

Procedure

All parents were contacted by letter to inform them of the nature of the research and request their assistance. Following their agreement to participate, detailed self-report questionnaire booklets were sent to the parents and collected by interviewers. Open-ended interviews were carried out in the parents' homes. Mothers and fathers were interviewed separately. The children were assessed in school and the completed checklist was obtained from the child's teacher. Teacher checklists were completed within two weeks and parental interviews carried out within one month of the child's assessment.

Measures

Academic attainments

The teacher questionnaires contained three checklists relating to the academic abilities of reading and use of written information, number skills and writing skills. As no standardised measures of these abilities were available which covered the early stages of pre-reading, number and writing skills as well as the more advanced skills, the checklists were constructed specifically for the study, based on those devised by Lorenz (1985) in a study of children with Down's syndrome. Checklist scores were obtained on all except the one child who was educated at home.

The reading and use of information checklist comprised 17 levels from "matches pictures of most common objects" to "reads with understanding to get information, e.g., newspapers, brochures". The number skills checklist comprised 22 levels from "discriminates between largest and smallest group of objects" to "does simple division sums" followed by "can do more advanced number work. Please specify . . .". The writing checklist comprised 19 levels from "holds pencil or crayon and attempts to scribble" to "writes imaginative pieces, creative writing". (See Appendix A.)

Each level within the checklists was scored as "Can't do" (1); "Can do with help" (2); or "Can do" (3). Separate scores were computed for the three different areas, and summed to create an Academic Attainments Index. The possible range of scores was 15 to 51 for reading, 22 to 66 for number, 19 to 57 for writing and 56 to 174 for the total index.

Details of means and distribution of scores for the three checklists and the Academic Attainments Index are given in Table 2. The skew of the distribution indicates that, in relation to the possible range of scores given above, the scores were grouped towards the lower end of the scales. Internal consistency reliability was assessed for the index and the three subscales, and found to be satisfactory. Coefficient alpha values (Cronbach, 1951) were 0.91, 0.93, 0.89 and 0.97 for reading, writing and number checklists and Academic Attainments Index, respectively.

TABLE 2

MEASURES OF CENTRAL TENDENCY AND DISTRIBUTION OF SCORES FOR THE THREE ACADEMIC ATTAINMENTS SCALES AND THE TOTAL INDEX (N = 117)

	Mean	Median	Min.	Max.	SD
Reading	28.15	26	17	50	7.95
Number	31.04	29	22	64	8.93
Writing	29.87	28	19	52	6.46
Academic Attainments Index	89.07	82	58	164	22.26

The scores on the three checklists were highly correlated: reading with number, r = 0.87, P<0.001; reading with writing r = 0.86, P<0.001; number with writing r = 0.89, P<0.001. It was therefore decided to use the total scores as the outcome measure. The validity of the measure was supported by the significant correlations between the checklist and index scores and the independent assessments of mental age on standardised tests: mental age with reading r = 0.74, number r = 0.73, writing r = 0.78, Academic Attainments Index r = 0.79, P<0.001 for all correlations.

In addition to the teacher checklists, reading attainment was assessed on the Spar Reading Test (Young, 1976) by the independent assessors. Only 20 children attained reading ages on this test. For these children the mean reading age was 6.8 years, range 5.9 to 9.7. The correlation of reading age with scores on the reading checklist was 0.72 (P<0.001) and with the scores on the Academic Attainment Index 0.70 (P<0.001). As the number attaining reading ages was small, only the wider measure of level of academic attainments, represented by the Academic Attainments Index scores, was used for further analysis. The relationship between the index scores and a number of descriptor variables was investigated.

Descriptor variables

These were mainly derived from the questionnaire which comprised a wide range of instruments of demonstrated reliability and validity (Sloper *et al.*, 1988). These measures were selected to reflect a broad spectrum of child and family characteristics which the literature suggests may be associated with child and family functioning. The variables used in the analysis were:

i) *Child variables*. Gender; chronological age; health problems; position in family; type of school (for the purposes of analysis classification was mainstream, unit, MLD, SLD: the three children in other placements were excluded); mental age and IQ; distractability and excitability based on the ADIECAS scale (Evans and Hogg, 1984) completed by teachers; behaviour problems — scores on the Behaviour Problems Index (Cunningham *et al.*, 1986) completed separately by mothers and fathers. The correlation between mothers' and fathers' scores was 0.77 (P<0.001); mothers' scores were used for the analysis to maximise numbers. Alpha reliability for the scale was 0.88. Amount of supervision and caretaking needed was assessed by mothers using the measure of Richman *et al.* (1982). In addition, where established instruments appropriate to this sample were not available, a number of measures were developed specifically for the study. These were:

a) A 13-item Communicative Style checklist completed by teachers. This measure was designed to assess the effective use of communication skills, both verbal and gestural, and was based on the work of Bates (1976) and Miller (1978). Scoring was on a three-point format (never, sometimes, often). Alpha reliability for the checklist was 0.92 and validity was indicated by a significant correlation with verbal age scores on the McCarthy Scales (r = 0.53 P<0.01).

b) An 11-item Appearance checklist completed by teachers. The scale was adapted from those of Richardson *et al.* (1985) and Flynn (1989) and concerned features relating to physical attractiveness (e.g., weight, posture) and physical features typically found in Down's syndrome (e.g., eye skin folds, tongue protrusion). Alpha reliability was 0.79, inter-rater reliability was assessed by completion of the items on the scale which did not require ratings of how often characteristics were present, by the researchers at the time of the children's assessment. The correlation between researchers' and teachers' scores was 0.69 (P<0.001). Teachers also rated the attractiveness of the child in relation to other children on an analogue scale. The correlation between this rating and Appearance checklist scores was 0.59 (P<0.001).

c) A 317-item Self-Sufficiency checklist completed separately by mothers and fathers. This scale was developed from the Pathways to Independence Scale (Jeffree and Cheseldine, 1982); additional items were included to increase sensitivity for the present sample, and scoring was on a three-point scale (can do, can do with help, cannot do) instead of yes/no, in order to overcome the problems of completion noted by Jeffree and Cheseldine for the original scale. Alpha reliability for the present study was 0.97, the correlation between mothers' and fathers' scores was 0.81 (P<0.001). Mothers' scores were used for the analyses as these were available for more children. Self-sufficiency scores showed a correlation of 0.71 (P<0.001) with the independent assessments of mental age on the McCarthy scale.

d) A measure of the frequency of the children's social contacts with other children was obtained through a questionnaire completed separately by mothers and fathers concerning the child's membership of organisations and informal play contacts. Scores were assigned on the basis of number and frequency of participation in these areas. Alpha reliability for the scale was 0.95, the mother/father correlation was 0.64 ($P<0.001$). Again mothers' scores were used in the analyses as they included more children.

ii) *Parental and family characteristics.* Social class; unemployment; parental educational level; type of housing; car ownership; marital status; parental status — natural, foster or adoptive; age of parent; number of children in family; financial problems, number and frequency of contact with various services. The former measures were based on parents' answers to questions in the self-report questionnaire. In addition, measures of type of housing and neighbourhood were rated by interviewers using ratings devised by Osborn *et al.* (1984). Published scales were used to address the following variables: personality — the Eysenck Personality Inventory (Eysenck and Eysenck, 1964); locus of control — the Brief Locus of Control Scale (Lumpkin, 1983); marital relationship — the Measure of Marital Satisfaction (Kelso *et al.*, 1984); amount of social support from the Inventory of Parent Experiences (Crnic and Greenberg, 1983); characteristics of the family system — the Family Enviroment Scale (Moos and Moos, 1981); sibling behaviour problems — Behaviour Checklist (Richman *et al.*, 1982) for siblings under 6 years of age, Rutter Scale A (Rutter *et al.*, 1970) for those aged 6 to 16; and interviewer ratings of parent-child relationships based on the method devised by Richman *et al.* (1982).

iii) *Life events.* Occurrence of life events in the last year was measured by a 42-item checklist (Cheang and Cooper, 1984). This checklist also measures the individual's perceptions of the strain of the events on a 10-point Likert-type scale.

iv) *Coping strategies.* The Ways of Coping Questionnaire (Revised) (Folkman and Lazarus, 1985) was adapted to measure the ways in which parents coped with problems concerning their children with Down's syndrome. Principle components factor analysis produced five factors: (1) Practical coping; (2) Wishful thinking; (3) Stoicism; (4) Seeking emotional social support; (5) Passive acceptance. Proportional coping scores for each factor were used in the analysis as these take into account individual differences in relative use of ways of coping (Vitaliano *et al.*, 1987).

For measures in (ii) to (iv) scores were obtained from both mothers and fathers. Details of all measures can be found in Sloper *et al.* (1988).

Statistical analysis

The relationships between the outcome variable of Academic Attainments Index scores and descriptor variables were examined using one-way analysis of variance for categorical descriptor variables and regression plots for continuous variables. The linearity of all relationships was examined to determine whether variables could be entered in a multiple regression analysis. All variables significantly related to outcome at the univariate level proved suitable for entry into multiple regression analyses.

Thus, a subset of descriptor variables was identified for inclusion in a multiple regression analysis. Descriptors were included if they were related to the outcome at the 5 per cent level of significance, or if there were theoretical grounds for inclusion.

Correlations between descriptor variables were computed and where variables were confounded (intercorrelations exceeded 0.59), bivariate analyses were used to determine which of the variables should be entered in the multivariable analysis.

RESULTS

Table 3 shows the variables significantly related to academic attainments scores in the univariate analyses.

Confounding of variables

Intercorrelations exceeding 0.59, and indicating statistical confounding, were found between a number of variables: mental age and IQ and self-sufficiency scores; Social

TABLE 3

Variables Significantly Associated with Academic Attainments in Univariate Analyses

	N	% of Variance associated	Significance level
(a) *Child*			
Mental age	121	+62.2	0.000
[Self-sufficiency	110	+35.7	0.000]
[IQ	121	+30.9	0.000]
Gender female	117	+11.7	0.000
Distractability	122	–9.6	0.001
Communicative style	122	+18.8	0.000
Chronological age	122	+13.9	0.000
Behaviour problems (M)	111	–7.8	0.003
Behaviour problems (F)	82	–11.3	0.002
Organisations Index	116	+9.6	0.001
[Social life index	100	+7.9	0.005]
(b) *Parental*			
Financial problems (M)	115	–7.9	0.002
Financial problems (F)	86	–5.9	0.024
Coping – wishful thinking (M)	104	–4.9	0.024
Coping – social support (M)	104	+4.4	0.032
Coping – social support (F)	79	+7.0	0.019
Locus of control (F)	86	+8.5	0.006
Family cohesion (F)	85	+15.0	0.000
Family expression (F)	83	+5.6	0.032
Intellectual-cultural orientation (F)	83	+15.6	0.000
Intellectual-cultural orientation (M)	106	+7.6	0.004
Active-recreational orientation (F)	84	+15.5	0.000
Active-recreational orientation (M)	107	+6.7	0.007
Moral-religious emphasis (F)	83	+7.0	0.016
Moral-religious emphasis (M)	105	+4.6	0.028
Mother-child relationship	119	+3.6	0.039
(c) *Socio-economic/demographic*			
Type of school	115	+31.4	0.000
Social class	117	+4.9	0.016
Father in employment	110	+4.2	0.031
Parental educational qualifications	120	+3.9	0.030
[School leaving age (F)	85	+5.7	0.028]
Mother in employment	117	+3.6	0.041
Car ownership	117	+5.6	0.010
Recent illness of mother	108	–4.1	0.036
Single parent family	118	–4.0	0.030
Mother's age	114	+3.5	0.047

Notes: +/– = direction of relationship: + denotes high scores on variable associated with high scores on Academic Attainments Index: – denotes high scores on variable associated with low scores on Academic Attainments Index.

(M) = measure derived from mothers' responses

(F) = measure derived from fathers' responses

[ ] = variable confounded with immediately preceding variable and not included in multivariable analysis

Life and Organisations Index scores; parental educational qualifications and fathers' school leaving age; and scores from mothers and fathers on variables which measured the same aspect of the child or family from both sources — the Behaviour Problems Index, financial problems and the three subscales of the Family Environment Scale of Intellectual-Cultural Orientation, Active Recreational Orientation and Moral Religious Emphasis.

In bivariate regression analyses mental age added significantly to the variance accounted for by IQ and self-sufficiency scores, and was also preferred as the most relevant to any explanation of the child's level of achievement at school. The Organisations Index score, which measured group membership and attendance, proved more powerful than the more general Social Life Index, which covered both organisations and friendship patterns, as the latter showed no significant association with scores on the Academic Attainments Index. Neither parental educational qualifications, measuring the possession of higher educational qualifications by either parent, nor fathers' school leaving age added significantly to the variance accounted for by each other. The former variable was selected for inclusion in the analysis as data were available on more subjects for this measure. In the case of confounded variables measured separately from responses of fathers and mothers, the analysis was run twice, once including mother variables and once including father variables, as it was deemed important to ascertain whether the views of either parent on child and family variables had any significant relationship with outcome. In the event, the same solution was obtained in both cases.

Results of multivariable analysis

Stepwise multiple regression analysis was computed using the Academic Attainments Index scores as the dependent variable and entering all variables significantly associated with these scores at the univariate level with the criterion for entry set at P=0.05.

Table 4 gives the results of the analysis. Mental age was found to be most strongly associated with academic attainments, and entered the equation first, accounting for 62 per cent of the variance. The majority of the variables showing univariate relationships with academic attainments lost significance after the entry of mental age. The other variables entering the equation therefore indicated the factors which contributed to the explanation of the variance over and above the children's basic cognitive development. Only type of school, gender, fathers' locus of control scores and chronological age of the child remained significant.

TABLE 4

STEPWISE MULTIPLE REGRESSION SOLUTION FOR FACTORS RELATED TO ACADEMIC ATTAINMENTS

Variable name	% variance associated	Addition to variance	Sig T in final equation
Mental age +	62.25	62.25	0.000
Type of school +	68.26	6.01	0.000
Gender female +	72.78	4.53	0.001
Fathers' locus of control +	75.78	3.00	0.004
Chronological age +	77.53	1.75	0.027

F = 45.55, df = 5,66, P = 0.000, N = 72, +/– = direction of relationship

The significance of type of school in the equation indicated that children in mainstream schools were likely to have the highest academic attainments, followed by those in units in mainstream schools, MLD schools and SLD schools, even after allowing for the difference in the mental ages of the children in the different types of schools. The entry of gender into the equation indicated that Academic Attainments Index scores were likely to be higher for girls than for boys, again after controlling for mental age differences, which also favoured the girls (Sloper *et al.*, 1988). Higher academic attainments were also associated with higher fathers' scores on the measure of locus of control, i.e., the father's view that events or outcomes were within his influence rather than external to it. The final variable to enter the equation was chronological age, with older children likely to have higher scores. In all, 78 per cent of the variance was accounted for by these factors.

DISCUSSION

The strong relationship between mental age and academic attainments in this group of children with Down's syndrome reflects the relationship found for non-handicapped children (e.g., Pellegrini *et al.*, 1987). Similarly, the finding that environmental factors were related to academic attainments after controlling for mental age has parallels with non-handicapped populations. These findings underline the importance of approaches which emphasise that the same factors which affect children without a biological handicap are also likely to affect children with Down's syndrome and need to be taken into account when planning intervention and education programmes for such children. As in studies of family adaptation in this cohort (e.g., Byrne *et al.*, 1988), the "normality" of the factors found to relate to outcome indicates that the biological condition of Down's syndrome is only one of a complex set of factors which influence child and family adaptation and outcome.

The first variable to enter the equation after mental age, type of school, indicated that children in mainstream schools were likely to have higher academic attainments than those in special schools, even after allowing for the difference in the mental ages of the children in different schools. Casey *et al.* (1988) showed similar findings for a smaller sample of children with Down's syndrome.

These findings have implications for both the educational placement and the curriculum for such children, and support the policy of placing children in mainstream schools wherever possible, as defined in the 1981 Education Act. Most of the children attending mainstream schools were in the higher end of the IQ range of the cohort at the time they entered school, and many received extra support in school. Even so, a number of children of equivalent IQ were placed in special schools. It is possible that those children had other problems at the time of admission to school, e.g., difficult behaviour. However, no evidence of differences in current problems could be found in the data: there were no significant differences between children in mainstream and special schools on measures of behaviour problems, self-sufficiency, health or family functioning, when mental age was controlled for.

The 21 children in mainstream schools and units attended 19 different schools in 11 different local education authorities. The amount of additional support they received varied considerably. Although teachers' training and experience in working with children with Down's syndrome or other children with learning difficulties was not formally measured in the study, the researchers' contacts with the teachers indicated that this varied widely. Some teachers had no previous experience or training whereas others had considerable experience and/or training. Unfortunately, the lack of formal measurement of these variables and the small numbers involved precluded investigation of the influence of such factors on the children's attainments. It is, however, an important aspect for further study.

The reason for the greater progress of children in mainstream schools in the traditional academic skills may lie in the different emphasis placed on the teaching of such skills in the different types of schools. Studies of teacher priorities (Bennett, 1976; de Souza and Bailey, 1981; Cleave *et al.*, 1982; Robson and Freeman, 1985) have found that teachers in mainstream primary schools were likely to place more emphasis on these academic skills, whilst those in special schools stressed self help, socialisation and language skills. This is not to say that the children in special schools were not taught academic skills, but that, as suggested by Lorenz *et al.* (1985), their lower attainment levels may be due to the differing importance placed on academic attainment. It may be supposed, therefore, that children attending special schools would show an advantage in terms of self-help and socialisation. However, findings from the cohort study suggest that type of school attended had little effect on self-help skills. Although self-sufficiency scores were significantly related to type of school, the relationship was accounted for by mental age, i.e., differences became non-significant when controlled for mental age. It appears, therefore, that the lesser emphasis on self-help and socialisation in mainstream schools did not have any disadvantageous effects on the children, and their attainments in this area were consistent with their mental age levels. The analyses of factors related to self-sufficiency indicated that family factors were more important than type of school attended (Sloper *et al.*, 1988).

The significant effect of gender indicated that, at this level, girls were likely to have higher academic attainments than boys. This has been demonstrated in other studies of children with Down's syndrome (e.g., Carr, 1988; Casey *et al.*, 1988) and in non-handicapped populations (e.g., Kellmer-Pringle *et al.*, 1966).

The fourth variable in the regression equation was fathers' Locus of Control scores. Higher academic attainments in the children were associated with internal locus of control, or the view that events and outcomes are within one's own influence rather than external to it. It is suggested that high scores on the Locus of Control scale reflect the person's ability to cope successfully through flexibility in the appraisal of his or her level of control in any specific situation (Fleming *et al.*, 1984). Internal locus of control has been shown in a number of studies to relate to higher academic achievement and higher aspirations, even after controlling for social class factors (e.g., Bar-Tal *et al.*, 1977, 1980; Maqsud, 1983).

The relationship between the fathers' Locus of Control scores and the children's academic attainments was significant even after controlling for type of school attended. It could not therefore be wholly due to a relationship between locus of control and parental influences on school placement. Why this characteristic of the fathers was related to the children's school achievement, while that of the mothers showed no relationship, is not clear, but mothers' and fathers' Locus of Control scores correlated only weakly in this sample ($r = 0.27$, $P=0.05$). It may be that the fathers' aspirations for the children were particularly pertinent to the children's academic attainments. Whether the relationship was due to fathers with higher Locus of Control scores being more actively involved in their children's education or to more diffuse factors related to the educational climate of the home cannot be ascertained from the analysis. Examination of the individual data on father's Locus of Control scores, along with the researchers' knowledge of the cohort, suggested that fathers who scored highly on this measure were those who took an active role with the child, instigated activities, took part in decision-making about the child, and took responsibility for the child equally with the mother. The low scorers, on the other hand, tended to be fathers who left decision-making and the main responsibility for the child to the mothers. Although they may have engaged in activities with the child, their involvement tended to be at the instigation of the mother. However, there was no significant association between fathers' Locus of Control scores and the quality of their relationships with the children. It is possible that fathers who were more directly involved with the children at their own instigation may have focused on such areas as academic and educational activities, reflecting a more "instrumental" than "expressive" role (Parsons and Bales, 1955; Gunz and Gubrium, 1972). In this case it seems likely that the association of fathers' Locus of Control scores with children's academic attainments reflected the effects of their involvement in their children's education as well as more diverse effects of family climate.

This implies that the role of fathers in the education of their children with Down's syndrome requires more attention. The main link between home and school is often through the mother, with less attendance at school events shown for fathers (e.g., Varlaam *et al.*, 1984). Although schools often find it difficult to involve fathers, policies which take fathers' involvement into account, and consider such factors as the time and content of school events, home visits and the active encouragement of fathers' participation, may be beneficial to the children.

Interestingly, the family background variables of social class and parental educational level were not significant factors in the final equation. Carr (1988) found that social class was related to academic achievements of young people with Down's syndrome after controlling for IQ. Similarly, this relationship has been shown for non-handicapped children. In the cohort both social class and parental educational qualifications have been found to be consistently related to the children's mental age scores in longitudinal data (Cunningham, 1986; Sloper *et al.*, 1988). However, in the present study the univariate relationships of these factors with academic attainments lost significance when controlled for mental age. Thus, there did not appear to be any residual effect of these particular family background variables over and above their effect on mental age scores.

The relationship of chronological age to academic attainments even after controlling for mental age suggests that the children gained in ability over time, presumably through greater quantity of instruction. It may also be related to the age at which schools introduced the children to reading, writing and number work, with some of the younger children having very little experience in these areas. Thus, their apparent delay may result from curriculum issues in the schools rather than child or family factors.

The teaching of the traditional academic skills to children with Down's syndrome has been a source of considerable debate (e.g. Gibson, 1978; Buckley, 1985; Carr, 1988). Given the importance placed on them in society, the increasing evidence that many children with mental handicap are able to attain at least basic levels in these skills, the contention that these skills "can be a source not only of pride but also of real pleasure to the young people concerned" (Carr, 1988, p. 425), and now the demands of the National Curriculum, further research on the particular aspects which contribute to greater progress would be useful. In particular, aspects of mainstream schooling, such as teacher expectations and training, curriculum, class size or peer interaction, which may relate to the children's academic progress need further study. The role of fathers in their children's education and the precise mechanisms whereby this may relate to academic attainment also have considerable implications both for school policies and for intervention with children with disabilities. Further research and evaluation needs to address the issue of how interventions on these aspects can produce improvements in functioning for children with Down's syndrome.

ACKNOWLEDGMENTS.—This research was supported by the Department of Health and Social Security, Grant no. 8ZB92. The authors would like to thank Ann Rangecroft and Garry Hornby for assistance in the research, Professor C. Kiernan for his comments on the manuscript, and all the parents, teachers and children involved in the study.

Correspondence and requests for reprints should be addressed to Dr P. Sloper, Hester Adrian Research Centre, University of Manchester, Manchester M13 9PL.

REFERENCES

Bar-Tal, D., and Bar-Zohar, Y. (1977). The relationship between perception of locus of control and academic achievement: review and some educational implications. *Contemp. educ. Psychol.*, 2, 181-199.

Bar-Tal, D., Kfir, D., Bar-Zohar, Y., and Chen, M. (1980). The relationship between locus of control and academic achievement, anxiety and level of aspiration. *Br. J educ. Psychol.*, 50, 53-60.

Bates, E. (1976). Pragmatics and sociolinguistics in child language. In Morehead, D. M., and Morehead, A. E. (Eds.), *Normal and Efficient Child Language*. Baltimore: University Park Press.

Bennett, N. (1976). *Teaching Styles and Pupil Progress*. London: Open Books.

Bodner-Johnson, B. (1986). The family environment and achievement of deaf students: a discriminant analysis. *Except. Children*, 52 (5), 443-449.

Buckley, S. (1985). Attaining basic educational skills: reading, writing and number. In Lane, D., and Stratford, B. (Eds.), *Current Approaches to Down's Syndrome*. London: Holt, Rinehart and Winston.

Buckley, S., and Sacks, B. (1987). *The Adolescent with Down's Syndrome: Life for the Teenager and for the Family*. Portsmouth: Portsmouth Polytechnic.

Byrne, E. A., Cunningham, C. C., and Sloper, P. (1988). *Families and Their Children with Down's Syndrome: One Feature in Common*. London: Routledge.

Carr, J. (1985). The development of intelligence. In Lane, D., and Stratford, B. (Eds.), *Current Approaches to Down's Syndrome*. London: Holt, Rinehart and Winston.

Carr, J. (1988). Six weeks to 21 years old: a longitudinal study of children with Down's syndrome and their families. *J. Child Psychol. Psychiat.*, 29 (4), 407-431.

Casey, W., Jones, D., Kugler, B., and Watkins, B. (1988). Integration of Down's syndrome children in the primary school: a longitudinal study of cognitive development and academic attainments. *Br. J. educ. Psychol.*, 58, 279-286.

Cheang, A., and Cooper, C. L. (1984). Psychological factors in breast cancer. *Stress Med.*, 1, 61-66.

Cleave, S., Jowett, S., and Bate, M. (1982). *And So To School· A Study of Continuity from Preschool to Infant School.* Windsor: NFER-Nelson.

Crnic, K. A., and Greenberg, M. T. (1983). *Inventory of Parent Experiences: Manual*. Seattle, Washington: University of Washington.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.

Cunningham, C. C. (1986). Early intervention: some findings from the Manchester cohort of children with Down's syndrome. In Bishop, M., Copley, M., and Porter, J. (Eds.), *Portage: More Than a Teaching Programme*. London: NFER-Nelson.

Cunningham, C. C., Sloper, P., Rangecroft, A., and Knussen, C. (1986). *The Effects of Early Intervention on the Occurrence and Nature of Behaviour Problems in Children with Down's Syndrome*. Hester Adrian Research Centre, University of Manchester. Final Report to DHSS.

DeSouza, D., and Bailey, T. J. (1981). What do parents, governors and professionals consider to be the most important aims for children with severe learning difficulties? *J. Br Inst. Ment. Handicap*, 8, 4, 130-131.

Duffen, L. (1979). For reading read listening. *Learning*, 1, 61-63.

Dunsdon, M. I., Carter, C. O., and Huntley, R. M. C. (1960). Upper end of range of intelligence in mongolism. *Lancet*, i, 565-568.

Evans, P. L. C., and Hogg, J. (1984). A classroom rating scale for use with mentally retarded children. *Br J. clin. Psychol.*, 27, 187-193.

Eysenck, H. J., and Eysenck, S. B. G. (1964). *Manual of the Eysenck Personality Inventory*. London: University of London Press.

Fleming, R., Baum, A., and Singer, J. E. (1984). Toward an integrative approach to the study of stress. *J. Person. soc. Psychol.*, 46, 939-949.

Flynn, M. C. (1989). *Independent Living for Adults with Mental Handicap: A Place of My Own.* London: Cassell.

Folkman, S. and Lazarus, R. S. (1985). If it changes it must be a process: a study of emotion and coping during three stages of a college examination. *J. Person. soc. Psychol.*, 48, 150-170.

Gibson, D. (1978). *Down's Syndrome: the Psychology of Mongolism*. London: Cambridge University Press.

Gumz, E. J., and Gubrium, J. F. (1972). Comparative parental perceptions of a mentally retarded child. *Am. J. ment. Defic.*, 77, 175-180.

Hewison, J. and Tizard, J. (1980). Parental involvement and reading attainment. *Br. J. educ. Psychol.*, 50, 205-215.

Jeffree, D. M. and Cheseldine, S. E. (1982). *Pathways to Independence.* Sevenoaks, Kent: Hodder and Stoughton Educational.

Kellmer-Pringle, M. L., Butler, N. R., and Davie, R. (1966). *11,000 Seven Year-olds.* London: Longman.

Kelso, J., Stewart, M. A., Bullers, L. and Eginton, R. (1984) The Measure of Marital Satisfaction: a questionnaire to screen parents for marital problems. *Child Psychiat. hum. Dev.*, 15 (2), 86-103.

Lorenz, S. (1985) Long term effects of early intervention in infants with Down's Syndrome. University of Manchester: unpublished Ph.D. thesis.

Lorenz, S., Sloper, P., and Cunningham, C. C. (1985). Reading and Down's syndrome. *Br. J. spec. Educ.*, 12, 2, 65-67.

Lumpkin, J. R. (1983). Validity of a brief locus of control scale for survey research. *Psychol. Reports*, 57, 655-659.

Marjoribanks, K. (1987). Ability and attitude correlates of academic achievement: family group differences.

J. educ. Psychol., 79, 2, 171-178.

Maqsud, M. (1983). Relationships of locus of control to self esteem, academic achievement and prediction of performance among Nigerian secondary school pupils. *Br. J. educ. Psychol.*, 53, 215-221.

Miller, L. (1978). Pragmatics and early childhood language disorders: communicative interactions in a half hour sample. *J. Speech hear. Disord.*, 43, 419-436.

Moos, R. H., and Moos, B. C. (1981), *Family Environment Scale: Manual.* Palo Alta, Calif.: Consulting Psychologists Press.

Nihira, K., Myers, C. E., and Mink, I. T. (1983). Reciprocal relationships between home environment and development of TMR adolescents. *Am J. ment. Defic.*, 88, 139-149.

Nihira, K., Mink, I. T., and Meyers, C. E. (1984). Salient dimensions of home environment relevant to child development. In Ellis, N. R., and Bray, N. W. (Eds.) *International Review of Research in Mental Retardation, Vol. 12*, Orlando, Florida: Academic Press.

Osborn, A. F., Butler, N. R., and Morris, A. C. (1984). *The Social Life of Britain's Five Year Olds.* London: Routledge and Kegan Paul.

Parsons, T., and Bales, R. F. (1955). *Family, Socialisation and Interaction Process.* Glencoe, I11.: Free Press.

Pellegrini, D., Masten, A. S., Garmezy, N., and Ferrarese, M. J. (1987). Correlates of social and academic competence in middle childhood. *J. Child Psychol. Psychiat*, 28, 5, 699-714.

Pieterse, M. and Treloar, R. (1981). *The Down's Syndrome Program.* MacQuarie University: Progress Report.

Richardson, S. A., Koller, H., and Katz, M. (1985). Appearance and mental retardation: some first steps in the development and application of a measure. *Am. J. ment. Defic.*, 89, 475-484.

Richman, N., Stevenson, J., and Graham, P. (1982). *Preschool to School — A Behavioural Study.* London: Academic Press.

Robson, C., and Freeman, A. (1985). Teachers' views on the relative importance of different curriculum areas for children with severe learning difficulties. *Res. Educ.*, 33, 57-62.

Rutter, M., Tizard, J., and Whitmore, K. (1970). *Education, Health and Behaviour.* London: Longman.

Shepperdson, B. (1988). *Growing Up with Down's Syndrome.* London: Cassell.

Sloper, P., Cunningham, C. C., Knussen, C., and Turner, S. (1988). *A Study of the Process of Adaptation in a Cohort of Children with Down's Syndrome and Their Families.* Hester Adrian Research Centre, Manchester University: Final Report to DHSS.

Varlaam, A., Woods, J., Mortimore, P., and Sammons, P. (1984). *Parents and Primary Schools.* London; ILEA Research and Statistics.

Vitaliano, P. P., Maiuro, R. D., Russo, J., and Becker, J. (1987). Raw versus relative scores in the assessment of coping strategies. *J. Behav. Med.*, 10, 1-18.

Young, D. (1976). *Spar Reading Test.* London: Hodder Stoughton Educational.

(*Manuscript received 28th November*, 1989)

APPENDIX

ACADEMIC ATTAINMENTS CHECKLIST ITEMS

Reading/use of written information

1. Matches pictures of most common objects.
2. Matches words of up to five letters.
3. Recognises his/her own name written down.
4. Recognises five to ten sight words.
5. Recognises and picks out labels, trade names etc. (e.g., brand names for soft drinks, chocolate bars, breakfast cereals).
6. Recognises up to 25 sight words.
7. Reads simple sentences.
8. Knows most letter sounds of the alphabet.
9. Reads simple books, e.g., early readers.
10. Recognises and acts appropriately to written signals, e.g., 'Danger', 'Bus stop', 'Exit'.
11. Builds simple words through knowledge of letter sounds.
12. Reads and acts appropriately to signs giving directions, e.g. street names, directories in shops.
13. Reads and follows a line of instructions, e.g., 'cut along dotted line'.
14. Uses complex phonics, e.g., digraphs, sound blends, etc.
15. Reads and follows a sequence of instructions, e.g., recipes, rules of a game.
16. Reads books, magazines for pleasure.
17. Reads with understanding to get information, e.g., newspapers, brochures.

Number

1. Discriminates between largest and smallest groups of objects.
2. Identifies groups with the same number of objects.
3. Makes a group of five objects.
4. Makes a group of nine objects.
5. When given a group of objects, adds one.
6. When given a group of objects, takes one away.
7. Names and matches symbols, 0 to 9.
8. Writes symbols, 0 to 9.
9. Writes symbols, 10 to 20.
10. Adds written numbers up to 10, with materials, (e.g., fingers, counting apparatus).
11. Adds written numbers up to 10, without materials.
12. Subtracts from written numbers up to 9, with materials.
13. Subtracts from written numbers up to 9, without materials.
14. Adds two numbers up to 20, with materials.
15. Adds two numbers up to 20, without materials.
16. Subtracts from numbers up to 20, with materials.

17. Subtracts from numbers up to 20, without materials.
18. Adds numbers, 20+, without materials.
19. Subtracts from numbers up to 50.
20. Does simple multiplication sums.
21. Does simple division sums.
22. Can do more advanced number work. Please specify ....

Writing

1. Holds pencil or crayon and attempts to scribble.
2. Scribbles purposefully with pencil or crayon.
3. Copies a circle.
4. Copies letters.
5. Copies Christian name.
6. Writes Christian name independently.
7. Writes Christian name and surname independently
8. Copies a simple sentence of four or more words.
9. Writes a simple sentence of four or more words.
10. Writes own name, address and telephone number.
11. Writes simple sentences, when dictated.
12. Writes more than one sentence on own.
13. Writes short notes, e.g., messages, shopping list.
14. Writes short personal letter.
15. Fills in printed forms, coupons.
16. Writes and addresses personal letters.
17. Writes short factual statements, e.g., accounts of events.
18. Writes short, descriptive passages.
19. Writes imaginative pieces, creative writing.

Br. J. educ. Psychol., 60, 299-311, 1990

THE EFFECTS OF ATTRIBUTION-BASED INTERVENTION AND STUDY STRATEGY TRAINING ON ACADEMIC ACHIEVEMENT IN COLLEGE FRESHMEN

BY FRANK VAN OVERWALLE AND MACHTELD DE METSENAERE
(Free University, Brussels, Belgium)

SUMMARY. The effectiveness of two brief remedial programmes on freshmen's academic performance was investigated. In the first study, a programme was developed to increase students' motivation for their study by using attribution change techniques. Participants (N=43) watched video-taped interviews with senior students relating the causes of their failures at the beginning of the first year, and how they had managed to improve their exam scores at the end of the year. The results showed that this attribution video manipulation significantly increased the number of students (18 per cent) who passed the final examinations at the end of the first year relative to control students who did not participate. In a second study, this attributional intervention was contrasted with another programme designed to enhance students' ability to remedy their learning difficulties by teaching appropriate study strategies. Participants (N=57) in this programme learned conventional rules of study strategy use, and applied these during short practice trials on existing course material. The attribution change programme again produced a significant increase (20 per cent) in the number of students who passed the final exams. In contrast, the learning strategy course had no effect on academic performance.

INTRODUCTION

DURING recent years, remedial intervention programmes at schools and universities aimed at helping students with learning difficulties have increasingly adopted attributional principles to improve performance after failure. Attribution-based programmes are derived from attributional conceptions of motivation (Weiner, 1985, 1986), learned helplessness (Seligman, 1975; Abramson *et al.*, 1978), and self-efficacy theory (Bandura 1977, 1982). These theoretical approaches assume that the detrimental consequences on a person's well-being, motivation and performance following failure are, in part, produced by the cause attributed for the negative outcome. For example, if a student attributes failure on an exam to low ability, then he or she will experience reduced hope and confidence in attaining success on the next exam because ability is a stable factor over which the student has little control. This, in turn, reduces active attempts to overcome the learning difficulty and leads to withdrawal from the task and decreased performance.

To overcome these negative consequences, attribution change programmes attempt to replace maladaptive causal beliefs, such as low ability, with explanations that are more temporary (unstable) and modifiable (controllable) by the student, such as insufficient effort or inappropriate study strategies (see Försterling, 1985; Van Overwalle *et al.*, 1989). This type of attribution increases the student's expectations for future success and encourages the belief that something can be done to overcome the learning obstacle. This in turn gives rise to increased persistence and performance (Weiner, 1986).

In a large number of studies, attributional principles have been successfully applied to produce improved performance and persistence in school tasks of students with learning difficulties (for an overview see Försterling, 1985). Most of these remedial programmes have been conducted with pupils from primary and secondary schools. For this young population, adaptive attributions such as effort are typically induced by providing appropriate verbal feedback about effort as the cause of success and failure. For example, after finishing a task, children are told that a wrong answer is caused by lack of effort, and that a right answer is due to high effort (see Försterling, 1985).

In this article, our main goal was to test whether attribution change techniques can also be used to improve the performance of freshmen at a university level. However, since university students have a longer history of academic successes and failures than younger students, it seems likely that their attributions are more stabilised. A simple causal feedback procedure, as is typically used with young children, may therefore be insufficient to change the maladaptive causal beliefs of college students. A study by Van Overwalle (1986), for example, demonstrated that simply telling freshmen that failure is caused by insufficient effort or inappropriate strategies did not improve their exam performance. Hence, in the present attribution programme, a procedure is adopted that was recently developed specifically for university freshmen by Van Overwalle *et al.* (1989) and Wilson and Linville (1982, 1985). In this training procedure, not only is the desired attribution provided, but also some antecedent information is given that precedes and determines causal judgment.

This manipulation of antecedent information is based on the attribution inference principles advanced by Kelley (1967, 1973) and Heider (1958). According to Kelley, an effect is attributed to a cause when there is covariation between that cause and the effect, that is, when cause and effect are always present or absent. Kelley proposed that different types of antecedent information can be used to detect cause-effect covariation, such as information about the past (i.e., consistency). For example, if a student failed a first exam but passed the second one (low consistency), we tend to infer that prior failure was due to causes that are unstable such as lack of effort or insufficient experience.

Another type of antecedent information used in attribution change programmes is based on Heider's notions of the finality of someone's behaviours, which promotes the belief of personal causation or control over the outcome of these behaviours. For example, imagine a student who passed an exam after studying for a long time, seeking extra information at the library, and asking for help from classmates. Since all these activities are directed at the unique goal of attaining high marks (equifinality), we tend to think that success was due to factors under the student's voluntary control, such as high effort and appropriate study approaches.

To induce adaptive causal beliefs for failure, these two types of antecedent information were manipulated in the present attribution programme. We showed our participants videotaped interviews with several second-year students who indicated that they had failed their midterm exams but had improved their scores on the finals at the end of the year (low consistency), and who also reported on their many attempts to remedy their failures and to find more adequate study techniques (equifinality). It was hypothesised that this information would induce the participants to attribute their failures to causes that are unstable and controllable. To reinforce this antecedent manipulation, the videotaped students also provided spontaneous explanations for their low outcomes in terms of insufficient effort, inappropriate study strategies or lack of experience, that are typically perceived as relatively unstable and strongly controllable (Van Overwalle, 1989).

Our hypotheses were tested in two experiments. In a first study, we examined the efficacy of attributional techniques for increasing freshmen's motivation and academic performance. This analysis was extended in a second study, where the attribution treatment was contrasted against as well as in combination with a conventional study strategy course designed to increase students' ability to remedy their learning problems. In both studies, the remedial treatments were provided after the students completed their midterm exams and were informed about their grades. Since Van Overwalle *et al.* (1989) suggested that focusing exclusively on students with very low grades may be less efficacious as little educational benefit can be gained by students who are too much at a disadvantage, students who failed and those who succeeded alike at the midterms were admitted to the programmes. To examine what pretest ability level is most suited for students to benefit from the remedial treatments, we analysed possible differences between low and high ability students after the experimental manipulation. The performance on the final exams (at the end of the year) served to measure the effect of the interventions.

EXPERIMENT 1

Method

The first study comprised two conditions. Students in an *attribution* condition watched videotaped interviews with second-year students relating the causes of their prior failures, while students in a *control* condition did not receive any treatment (but followed the same academic courses and took the same exams). Both these conditions comprised students who had either failed or passed the midterm exam.

Sample. Participants were 118 freshmen enrolled in an economics or production management engineering study option at the Faculty of Economics of the Vrije Universiteit Brussel (Belgium). Two students in the control condition could not be identified and were excluded, leaving 116 participants. The most important course in the first year of both study options is "Introductory Economics". The attributional treatment took place during regular tutorial hours after the midterm exam as part of the course, and all students present participated.

Participants were randomly assigned to the two conditions, using existing class groups (formed in alphabetical order) to minimise interferences with the regular organisation of the course. There were two such class groups in each study option. Within each option, the classes followed the same curriculum, given by the same instructors using the same format (only teaching assistants differed sometimes). Moreover, the curricula of the two options overlapped to a large extent. The classes were randomly assigned to one of the two conditions, with the provision that both study options were represented in each condition. The number of students in the attribution condition was 43, and in the control condition, 73.

Attribution video manipulation. The attributional manipulation took place after the first hour of regular tutorial classes for the economics course and lasted about 50 minutes. The regular teaching assistant was present during this period, but was instructed not to participate.

The experimenter presented himself as a researcher looking into the causes of success and failure in the freshman year. He explained that research had revealed that a large number of freshmen experience difficulties in their studies. But, instead of reciting dry statistical results, he preferred to discuss with the students what they thought about the causes of their outcomes on the midterm economics exam. Each student then listed the causes for his or her midterm outcome, and the answers were written on the blackboard.

Next, the experimenter announced that he would show some real-life experiences of other, older students during their first year. He explained that he had interviewed several students in the upper class years who had experienced some difficulties in their first year, and had asked them to list the reasons for their failure and to explain how they had overcome their learning problems. Five of these interviews would be shown on video. Next, a psychology professor would be seen briefly discussing the results of a previous strategy course aiming at alleviating the learning problems of freshmen. The students were instructed to pay careful attention to the information because they would later be questioned about it.

The tapes were of actual interviews with three male and two female economics students. To induce unstable and controllable attributions, these students were specifically selected so that the following causal antecedent information was provided in the course of the interviews. First, all the interviewees reported that they had obtained low scores on their initial midterm economics exams but had improved these scores at the subsequent final exam, reflecting low consistency of initial failure. Second, they also reported on their many attempts to remedy their failures by trying out various alternative study approaches, implying equifinality for success obtained on the second exam.

In addition to the information provided on causal antecedents, the interviewees also reported on the causes of their failures: lack of study effort and inadequate revision of the subject matter (student #1); insufficient preparation of problems given in the tutorials (student #2); time needed to adapt to life and work at university and lack of co-operation with fellow students (student #3); lack of experience at the first exams, lack of a study technique for structuring and summarising the subject matter (student #4); and lack of an efficient method based on theoretical understanding for solving the tutorials (student #5). These attributions all fall in the rubric of insufficient effort, lack of experience, and ineffective strategies that are unstable and controllable. The information was authentic, thus resulting in a believable presentation. To increase the clarity of the presentation, the students' accounts were briefly rehearsed before recording and the interviews were not interrupted by questions from the interviewer during the final take. The interviews lasted 13 minutes.

To increase the validity of the interviewees' accounts of study problems and to underscore the necessity of active effort in trying out alternative study approaches, a brief (2 minutes) videotaped explanation by a psychology professor was then shown. In the videotape, the professor discussed a remedial programme that had been devised to assist freshmen with learning problems by providing them with feedback about the inadequacies of their study strategies. He explained that only participants who had attempted to improve their study approaches showed an increase in their academic scores. He illustrated these results with a figure indicating that students who had learned and implemented new study strategies increased exam scores by 3 marks out of 20, whereas those who had not done so showed no changes (these data were authentic, Van Overwalle, 1986). In his talk, the professor did not refer to any particular study strategy or approach to explain these improvements. Rather, students' personal responsibility and effort in solving their study problems was stressed.

After watching the videotape, the participants were asked to write down what they considered to be the most important information provided during the interviews and their personal reactions to this information, in order to intensify further cognitive work on the videotapes shown. Next, during a short discussion, the students reported on their written responses. These most often involved their perceptions of the important causes of failure at the midterms and their prescriptions for obtaining success on subsequent exams. Again, these attributions were written on the board. To conclude the discussion, the experimenter emphasised that seeking the causes for learning difficulties was the first crucial step towards the resolution of these problems. In this respect, the experience of other students was stressed as a valuable input in thinking about personal study problems. The experimenter underscored this learning process by pointing to the larger number of causes that the students had reported during the last discussion, as compared to the causes given before the videotape was shown.

Performance measurements. The outcome on the economics midterm exam held after the first semester (February) served as a pretest for the present study and was used as covariate. The final exam of the economics course, covering the material of the second semester, was held at the end of the second semester (May), and served as post-test for the treatment. Both exams included a theory (focusing on economic theory), a tutorial (focusing on the application of these economic principles), and a current issues section (focusing on students' readings of current economic topics).

To assess possible generalisation effects on overall academic performance, additional outcome measurements were gathered on the final exams at the end of the academic year (during the exam session in June-July). The final exams include all courses given during the freshmen year and usually cover all material from the beginning of the course. The exams were scored by the regular instructor.

The first generalised measure was the passing rate, that is, the number of students in each condition who passed the exam session. The second measure was the average score at all final exams. Other generalised measures included specific information regarding failure on the finals. When a student fails a final exam at the Free University of Brussels, there are two possibilities. A score between 7 and 11 out of 20 denotes failure on that exam (i.e., failure mark), but the student can still pass the final exam session if the average score for all exams is above 11. However, a student cannot pass if a score on any exam drops below 7 (i.e., no-passing mark). For all final exams, we tallied whether students had received either of these two types of marks.

Results and Discussion

Performance effects. We conducted a series of 2 (attribution vs. control) analyses of covariance with several post-training measurements as dependent variables and respective pretest scores as covariates. Means and standard deviations of the performances after the intervention are shown in Table 1. As predicted, significant differences were found between attribution and control conditions.

TABLE 1

MEANS (AND STANDARD DEVIATIONS) OF THE PERFORMANCES AFTER THE TREATMENT: EXPERIMENT 1

	Control (N = 73)	Attribution (N = 43)
Post-test Exam		
Theory[a]	10.64	11.18
	(3.66)	(4.03)
Tutorials[b]	9.31	8.89
	(4.65)	(4.57)
Current issues[c]	12.83	12.97
	(3.81)	(3.40)
Final Exams[d]		
Passing rate	0.44	0.64
	(0.50)	(0.49)
Percentage passed	45.20	62.79
Average mark	11.02	11.43
	(3.41)	(3.69)
Failure marks	0.66	0.65
	(0.48)	(0.48)
No-passing marks	0.53	0.32
	(0.50)	(0.47)

NOTE: Cell entries are post-test means with standard deviations in parentheses. Maximum score for exams = 20. Passing rate, Failure marks and No-passing marks are coded with 0=no and 1=yes. The means are adjusted for pretest scores on [a]theory, [b]tutorials, [c]current issues, [d]all three pretest scores.

The analyses revealed no differences for the theory, tutorial and current issues scores on the economics exam held after the treatment. (All significance tests reported in this article are at the 0.05 level.) However, there was a significant increase in the number of students who passed the final exam session, F (1, 111) = 5.37, $P<0.03$. The size of this effect is illustrated when we look at the proportion of students who passed in each condition. Sixty-three per cent of the attribution students passed, whereas only 45 per cent of the control students did so. This difference approached significance, χ^2 (1, N=116) = 3.35, $P<0.07$. Further analysis indicated that the improved passing rate was mainly due to a significant difference in no-passing marks, F (1, 111) = 5.92, $P<0.02$. Participants in the attribution condition received fewer no-passing marks (adjusted M = 0.32, SD = 0.47) than in the control condition (M = 0.53, SD = 0.50). In contrast, the average and failure marks did not differ between conditions.

Taken together, these findings indicate that students who watched the attribution video did not systematically improve their scores on the economics post-test, nor on the final exams. However, many of them successfully avoided extreme low (no-passing) scores at the finals so that they could pass. These results are less contradictory than it may seem. Since many experimental participants succeeded at the economics midterms, they probably did not experience much difficulty with this course, and perhaps therefore applied what they learned from the videotaped accounts on other material that they perceived as more problematic, with the consequence that generalised no-passing grades were improved rather than economics scores.

Effect of pretreatment ability. To investigate whether pretest ability had any differential impact on the attribution change treatment, we divided the sample into three equal groups of low, moderate and high ability, based on students' mean scores at the economics midterm (i.e., theory, tutorials and current issues). Midterm scores of 11.1 and 13.4 (out of 20) divided the three ability levels. Post-test and generalised performance following the experimental treatment was then analysed with a 2 (attribution vs. control) x 3 (ability level) analysis of covariance. If pretest ability makes a substantial difference on treatment outcomes, interaction effects between ability level and treatment conditions should emerge. However, none of the interactions reached significance, indicating that the attribution intervention was equally effective for all participants regardless of their initial pretest levels. Thus, the intervention improved no-passing marks and passing rate for low, moderate and high ability students alike.

EXPERIMENT 2

The results from the first study clearly demonstrate that attributional change techniques are effective tools to enhance students' *motivation* to study. In this second experiment, our aim was to extend this motivational approach to treatment programmes that focus on freshmen's learning skills and *abilities*. Therefore, we compared the effects of the attributional change programme with a brief study strategy course that was designed to increase students' ability to remedy their learning difficulties.

In this strategy course, students were directly instructed on conventional techniques and strategies for managing their study time and for studying prose text. The course was based on study training textbooks devised by van Hout-Wolters, Jongpier and Pilot (1983) and Van Wanzeele (1982), and on a similar training programme developed by Dansereau *et al.* (1979). We assumed that most college students have basic reading and problem solving skills. Therefore, the course focused on the *when* and *why* of learning techniques in order to foster the students' awareness in selecting and utilising strategies, rather than emphasising careful instruction of the rules for proper strategy execution.

Three main topics were addressed which incorporate several strategies known to facilitate learning. The first topic dealt with time management and formulation of attainable study goals (Bandura and Schunk, 1981; Morgan, 1985, 1987). The second topic covered various techniques that enhance the comprehension of prose text such as underlining and note-taking to identify main ideas of the text (Brown and Smiley, 1977; Annis, 1985; Stevens, 1988). The third topic dealt with techniques that improve the cognitive organisation and retention of prose content such as skeleton outlines, flow charts and similar aids to summarise and logically organise the material (Holey *et al.*, 1979; Taylor, 1982; Geva, 1983; Taylor and Beach, 1984), and cumulative rehearsal and self-testing to increase and monitor one's retention (André and Anderson, 1979; Wong, 1985; Schommer and Surber, 1986).

We predicted that the study strategy course as well as the attribution programme would improve freshmen's performance on subsequent examinations. This prediction was based on the assumption that the attributional testimonies from the videotape would enhance the students' motivation to study, whereas the teaching of appropriate study strategies would increase the students' ability to deal with their study difficulties. Both motivational and cognitive components were presumed to facilitate learning and improve performance. We made no predictions as to which of the two treatments would be most effective, and one of the purposes of this second study was to examine possible differences in their effectiveness. Further, assuming that the attribution and the strategy treatment focus on two different components of the learning process, we anticipated that their combination would produce the highest gain in performance.

Method

Four different conditions were provided in this study. In an *attribution* condition, participants watched videotaped interviews with senior class students; in a *strategy* condition, they followed a study strategy course; in a *combined* condition, they followed both the attribution and the strategy programmes (in this order); and students in the *control* condition did not receive any treatment, but followed the same academic courses and took the same exams as the other experimental participants.

Sample. Participants were 124 freshmen enrolled in the first year at the law department of the Vrije Universiteit Brussel (Belgium). The experiment was conducted during discussion hours of the "History of Belgian Politics" course, and all students present in class participated. Again, existing class groups (divided by alphabetical order) were used to avoid interference with regular teaching. All classes followed the same curriculum given by the same instructors using the same format (only teaching assistants sometimes differed). Eight out of ten existing class groups were randomly assigned to one of the four conditions. The number of participants in the attribution condition was 35, in the strategy condition 31, in the combined attribution and strategy condition 26, and in the control condition 32.

Attribution video condition. The attributional treatment was similar to that in Study 1, save for the actual content of the videotape. Three female and two male senior law students related their study difficulties and failures in the first year. They also reported on the causes of their failures: lack of appropriate time management (student #1); too much emphasis on the chronology rather than on the main themes in the text (students #2, 4 and 5); lack of assistance from others (student #2); lack of elaboration and organisation in the exam answers (student #3); too much rote learning (student #5).

After the interviews (16 minutes), a psychology professor briefly (2 minutes) discussed a fictitious research programme that investigated students' spontaneous use of study techniques. As in the first study, the professor explained that study strategies were important to attaining success at the exams. Apart from the videotape, all other aspects of the attribution video intervention were identical to those in Study 1.

Study strategy condition. The study strategy course was provided during class hours and lasted about 60 minutes. Each strategy addressed in the training course was presented using the same format. First, the trainer introduced the strategy by explaining for which sort of learning difficulties it would be most helpful. Next, he briefly presented the basic rules for applying the strategy, illustrating its use by drawing several examples from the students' curriculum. The participants were then given the opportunity to practise the technique using exercises from their course material. Finally, a short discussion followed in which the participants could express their own experiences and views, and listened to their peers and the trainer who gave feedback and guidance. The course consisted of three sections addressing strategies for time management, reading and studying from textbooks.

In the first section on time management, the participants learned how to plan and set attainable study goals, allocate time within a schedule, and monitor their activities. They were also instructed on how to act when the allotted time was exceeded. Model examples of time schedules were provided to illustrate these basic rules.

In the second section on reading from textbooks, the participants were told that they might not be able to remember everything in the text. Therefore, they should analyse the major goals of the course and accordingly identify the main ideas and important supporting details by underlining or summarising them while reading the textbook.

In the third section on studying the course material, they learned how to utilise skeleton outlines to represent the main ideas that they had identified and their logical or textual relationships. Similar structural aids such as flow charts, listings of key terms and the like were also presented. It was explained that building a macro-structure of the text can help to better organise, integrate and memorise the subject material in less time than that involved in simply rereading the text. Finally, it was demonstrated how such structural aids can be used to rehearse the main themes and an increasing number of important details and how they can help to formulate questions for controlling the accuracy of one's retention.

Combined attribution and strategy condition. The combined treatment consisted of the attributional video intervention, immediately followed by the study strategy course. The total duration was 110 minutes.

Performance measurements. Each year, the law department organises one midterm exam, alternately on the course "History of Law" and on "History of Belgian Politics". Both are major courses in the law curriculum of the freshmen year, and both involve historical analyses of legal and political developments. When the present study was conducted, the midterm exam (February) was held on "History of Law", which served as pretest. The video-taped interviews, however, reported on the exam from the year before, "History of Belgian Politics", and the strategy training was also mainly focused on this course. Hence, the final exam on this latter course (June) served as post-test. The generalised performance measures were identical to those in Study 1.

Results and Discussion

Performance effects. To analyse the data, we conducted a series of 2 (attribution) x 2 (strategy) analyses of covariance with midterm outcome as covariate and the various post-treatment measurements as dependent variables. The means and standard deviations are shown in Table 2.

With respect to the attribution programme, the same pattern of positive results was found as in Study 1. There was a significant increase in the passing rate at the final exam session, $F(1, 119) = 6.32$, $P<0.02$. Again, this improvement was mainly due to a decrease

TABLE 2

Means (and Standard Deviations) of the Performances after the Treatment: Experiment 2

	TREATMENT			
	Control (N = 32)	Attribution (N = 35)	Strategy (N = 31)	Attribution and Strategy (N = 26)
Post-test Exam	7.73	8.28	7.13	7.57
	(4.25)	(4.14)	(4.06)	(4.63)
Final Exams				
Passing rate	0.32	0.56	0.25	0.42
	(0.49)	(0.50)	(0.49)	(0.50)
Percentage passed[a]	20.00	37.93	10.00	31.58
Average mark	9.94	10.78	9.47	9.63
	(2.69)	(3.33)	(2.70)	(3.96)
Failure marks	0.80	0.75	0.88	0.79
	(0.42)	(0.40)	(0.38)	(0.40)
No-passing marks	0.55	0.37	0.70	0.54
	(0.51)	(0.51)	(0.50)	(0.50)

NOTE: Cell entries are post-test means with standard deviations in parentheses. Maximum score for exams = 20. Passing rate, Failure marks and No-passing marks are coded with 0=no and 1=yes. All means are adjusted for the mid term pretest scores. [a]Percentage of participants with pretest scores below 11 (see text).

in no-passing marks, F (1, 119) = 4.41, $P<0.04$, since the average mark and failure marks did not differ significantly between conditions (although the means pointed in the expected direction). No other significant main or interaction effects were found.

Although the results indicate that more students passed the exam session after participating in the attribution programme, the passing rate — expressed in percentages — nonetheless showed no clear differences between conditions. This is due to the fact that percentages cannot be adjusted for pretest differences. Indeed, the participants in the attribution condition obtained lower pretest scores than in the other conditions (e.g., 5.44 in the attribution and 8.94 in the control condition), and this concealed true treatment effects. We found that the low pretest performance in the attribution condition was mainly due to a smaller number of students with marks above 11 out of 20. Hence, we excluded all students with pretest scores above 11, and recalculated the percentages. After this adjustment, we found that 35 per cent of all students who had watched the video passed the final session, whereas only 15 per cent of the other students did so. This difference was statistically significant, χ^2 (1, N=88) = 4.71, $P<0.03$.

With respect to the strategy programme, we did not find any main or interaction effects. Thus, contrary to our expectations, the study strategy course did not increase participants' ability to remedy their learning problems.

Effects of pretreatment ability. As improvements were only found after the attribution treatment, we concentrated further on this experimental factor. (We also analysed the strategy course but found no systematic differences between ability levels). Again, the entire sample was divided into three equal groups of low, moderate and high ability based on students' pretest midterm scores. Scores of 5.0 and 9.5 (out of 20) divided the three pretest ability levels.

Post-test and generalised performance measures were analysed in a 2 (attribution) x 3 (ability level) analysis of covariance. We found significant interactions for passing rate, F (2, 117) = 5.93, P<0.01; no-passing marks, F (2, 117) = 7.55, P<0.001; and the interaction for the grade average at the finals also approached significance, F (2, 117) = 2.76, P<0.07. Simple main tests indicated that these interaction effects were due to highly significant performance differences between the attribution and the other conditions for students of moderate ability: they attained higher passing rates (adjusted M = 0.76, SD = 0.44 vs. M = 0.14, SD = 0.35, F (1, 36) = 23.85, P<0.001), fewer no-passing grades (M = 0.16, SD = 0.39 vs. M = 0.74, SD = 0.46, F (1, 36) = 18.18, P<0.001), and higher grade averages at the finals (M = 11.71, SD = 1.80 vs. M = 9.61, SD = 1.75, F (1, 36) = 13.13, P<0.001). In contrast, students of low and high ability level did not show any academic gains following the attribution treatment. Thus, the attribution manipulation was especially beneficial for participants with moderate ability. Perhaps many participants with low pretest scores were too disappointed to be helped by the attribution treatment, whereas a great number of participants with high scores were quite satisfied with their grades and found it unnecessary to increase them any further.

The finding in this second study that moderately talented participants benefited most from the attribution manipulation stands in contrast with the first study, where the same treatment proved equally effective across all ability levels. A possible explanation lies in the fact that the variance of pretest grades in Study 1 was much smaller (SD = 3.01) than in Study 2 (SD = 4.63). Hence, pretest ability differences may simply have been too small with the consequence that they had no effect on treatment outcomes in the first attribution change programme.

DISCUSSION

Attribution change

The beneficial effect of the present attribution video manipulation contributes to a growing literature that attests to the effectiveness of attributional techniques in promoting the learning outcomes of freshmen students (Wilson and Linville, 1982, 1985; Van Overwalle *et al.*, 1989). The present study is, however, also important in its own right, since one of the most critical parameters of academic achievement in European universities, the passing rate, was significantly enhanced by the attribution treatment. In both experiments, the number of students who passed the finals increased by about 20 per cent after watching the attribution video manipulation. Prior attribution experiments have not always been able to attain such generalised performance effects. The video manipulation in these experiments seems particularly interesting because the procedure is straightforward and short, and can easily be applied in class.

The facilitative effects of the attributional treatment were observed for freshmen in the economics as well as the law department. The present videotape procedure thus seems effective over a broad spectrum of the academic curriculum provided that the material used in the manipulation is relevant to the particular study option. Another encouraging finding is that the increased passing rate was mainly produced by a substantial drop of very low (no-passing) grades on one or more exams that otherwise would have prohibited students from passing the exams. This suggests that the attribution manipulation was most effective for freshmen who were highly at risk of failure, an interpretation that is further supported by the finding that gains were most evident among students with moderate pretest scores (see experiment 2). However, it would be premature to infer that candidates for an attribution treatment can be selected *a priori* on the basis of such pretest grades. Indeed, under some conditions (e.g., when differences between pretest scores are relatively small, see experiment 1), low and high ability students may benefit from an attribution programme as well.

Since self-report data were not available in this study, it is not immediately evident which changes in participants' beliefs contributed to their academic improvements. We cannot completely rule out alternative explanations that are not based on changes in attributional thinking. For example, it might be argued that students recognised that they were in an experimental group and that the experimenter expected them to improve their grades with the consequence that this expectation led to enhanced performance (cf. Hawthorne effect). However, such an explanation seems implausible. If a Hawthorne effect had taken place, it is unclear why it affected only the attribution and not the strategy condition, both conducted by the same experimenter who held the expectation that both conditions would be effective. Moreover, prior research by Van Overwalle (1986) demonstrated that mere verbal persuasion to change freshmen's attributions does not result in any gains, whereas laboratory experiments by Wilson and Linville (1982, 1985) showed enhanced performance following attribution treatments as compared to carefully matched placebo control conditions, which effectively rule out experimental artefacts. Taken together, it can be concluded that the improvements observed in the present studies result from the inducement of attribution change. Nonetheless, measures of students' causal thoughts or placebo control conditions are clearly needed to validate this assertion firmly, and therefore are a priority in our current research.

Study strategies

Teaching *when* and *why* powerful study techniques should be applied did not render learning more effective. This is a puzzling result, because there is a vast body of research showing that effective study techniques facilitate learning and retention of the course material. Why then did the present strategy course fail to produce such effects?

A possible explanation is that we incorrectly assumed that most freshmen have basic study skills upon entering university. Consequently, just telling students during a one hour course to underline main ideas in the text or to formulate questions for testing their retention may be insufficient to produce more accurate recall of important elements. However, this explanation seems unlikely. Although careful training of techniques can indeed facilitate learning, previous research indicates that a simple explanation of the procedure is often sufficient for older students to identify main themes, direct their study time to these elements and to generate adequate questions (see Brown and Smiley, 1978; André and Anderson, 1979).

A more plausible explanation for the low efficacy of the strategy programme is that many students often do not realise that their current learning strategies are less effective than the newly taught procedures. As a consequence, they are not motivated conscientiously to maintain and follow these procedures after instruction has ended. Various approaches have been suggested to extend the duration of strategy use. Most of these involve an attempt to increase the students' metacognitive awareness of the advantages produced by powerful strategies. Pressley *et al.* (1988), for example, successfully maintained strategy use by giving pupils the opportunity to compare the use of a new strategy with their regular or other less effective learning methods.

Another approach to stimulate long-term use of learning strategies involves *strategy-specific* attributions, that is, additional instruction that success and failure are dependent on the active use of the strategy taught during the training programme. In two studies with school children, for instance, teaching such strategy-specific attributions together with conventional instruction on strategy procedures produced more durable performance gains than the training of strategies alone (Reid and Borkowski, 1987; Borkowski *et al*, 1988). However, in the present study, the combination of an attribution and a strategy treatment did not result in substantially greater performance increases, presumably because the communicated attributions did not focus on the specific and technical aspects of learning strategies. Rather, the videotaped accounts dealt with general causes of academic success and failure (before any specific strategy techniques were taught). In order to alter the perceived usefulness of study techniques, perhaps attributions must be specifically tied to them. Thus, the inclusion of a specific attributional component in the training of study strategies may resolve some of the dilemmas of conventional strategy courses. How to integrate attribution change with strategy training within a freshmen population, however, remains a question for future research.

Acknowledgements.—We are grateful to B. Weiner, J. Juvonen, and R. Govender for their valuable comments on an earlier version of this manuscript.

Correspondence and requests for reprints should be addressed to Mr Frank Van Overalle, Free University Brussels, Department of Psychology, Pleinlaan 2, B-1050, Brussels, Belgium.

REFERENCES

Abramson, L. V., Seligman, M. E. P., and Teasdale, J. D. (1978). Learned helplessness in humans: critique and reformulation. *J. abnorm. Psychol.*, 87, 49-74.

André, M., and Anderson, T. H. (1979). The development and evaluation of a self-questioning study technique. *Reading Res. Q.*, 14, 605-623.

Annis, L. F. (1985). Student-generated paragraph summaries and the information-processing theory of prose learning. *J. exp. Educ*, 54, 4-10.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioural change. *Psychol. Rev*, 84, 191-215.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *Am Psychol.*, 37, 122-147.

Bandura, A., and Schunk, D. H. (1981). Cultivating competence, self-efficacy and intrinsic interest through proximal self-motivation. *J. Person. soc. Psychol.*, 41, 586-598.

Borkowski, J. G., Weyhing, R. S., and Carr, M. (1988). Effects of attribution retraining on strategy-based reading comprehension in learning-disabled students. *J. educ. Psychol.*, 80, 46-53.

Brown, A. L., and Smiley, S. S. (1978). The development of strategies for studying texts. *Child Dev.*, 49, 1076-1088.

Dansereau, D. F., Collins, K. W., McDonald, B. A., Holley, C. C. D., Garland, J., Diekhoff, G., and Evans, S. H. (1979). Development and evaluation of a learning strategy training program. *J. educ. Psychol.*, 71, 64-73.

Försterling, F. (1985). Attributional retraining: a review. *Psychol. Bull.*, 98, 495-512.

Geva, E. (1983). Facilitating reading comprehension through flowcharting. *Reading Res Q*, 18, 384-405.

Heider, F. (1958). *The Psychology of Interpersonal Relations.* New York: Wiley.

Holey, C. D., Dansereau, D. F., McDonald, B. A., Garland, J. C., and Collins, K. W. (1979). Evaluation of a hierarchical mapping technique as an aid to prose processing. *Contemp. educ. Psychol*, 4, 227-237.

Kelley, H. H. (1967). Attribution in social psychology. *Nebraska Symposium on Motivation*, 15, 192-238.

Kelley, H. H. (1973). The process of causal attribution. *Am. Psychol.* 28, 107-128.

Morgan, M. (1985). Self-monitoring of attained subgoals in private study. *J. educ. Psychol.*, 77, 623-630.

Morgan, M. (1987). Self-monitoring and goal setting in private study. *Contemp. educ. Psychol.*, 12, 1-6.

Pressley, M., Levin, J. R., and Ghatala, E. (1988). Strategy-comparison opportunities promote long-term strategy use. *Contemp. educ. Psychol.*, 13, 157-168.

Reid, M. K., and Borkowski, J. G. (1987). Causal attributions of hyperactive children: implications for teaching strategies and self-control. *J. educ. Psychol.*, 79, 296-307.

Schommer, M., and Surber, J. R. (1986). Comprehension-monitoring failure in skilled adult readers. *J. educ. Psychol.*, 78, 353-357.

Seligman, M. (1975). *Helplessness: On Depression, Development and Death.* San Fransisco: Freeman.

Stevens, R. J. (1988). Effects of strategy training on the identification of the main idea of expository passages. *J. educ. Psychol.*, 80, 21-26.

Taylor, B. M. (1982). Text structure and children's comprehension and memory for expository material. *J. educ. Psychol*, 74, 323-340.

Taylor, B. M., and Beach, R. W. (1984). The effects of text structure instruction on middle-grade students' comprehension and production of expository text. *Reading Res. Q.*, 19, 134-146.

van Hout-Wolters, B., Jongpier, P., and Pilot, A. (1983). *Studiemethoden.* Utrecht, The Netherlands: Spectrum.

Van Overwalle, F. (1986). Verhogen van studieprestaties bij eerstejaars door attributie retraining. [Improving performance of freshmen through attribution retraining] *Tydschrift voor Onderwijsresearch*, 11, 225-240.

Van Overwalle, F. (1989). Structure of freshmen's causal attributions for exam performance. *J. educ. Psychol.*, 3, 400-407.

Van Overwalle, F., Segebarth, K., and Goldchstein, M. (1989). Improving performance of freshmen through attributional testimonies from fellow students. *Br. J. educ. Psychol.*, 59, 75-85.

Van Wanzeele, L. (1982). Gids studiemethode. Unpublished manuscript, Free University Brussels, Belgium.

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychol. Rev.*, 92, 548-573.

Weiner, B. (1986). *An Attributional Theory of Motivation and Emotion.* New York: Springer-Verlag.

WILSON, T. D., and LINVILLE, P. W. (1982). Improving the academic performance of college freshmen: attribution therapy revisited. *J. Person. soc. Psychol.*, 42, 367-376.

WILSON, T. D., and LINVILLE, P. W. (1985). Improving the performance of college freshmen with attributional techniques. *J. Person. soc. Psychol.*, 49, 287-293.

WONG, B. Y. L. (1985). Self-questioning instructional research: a review. *Rev. educ. Res.*, 55, 227-268.

(*Manuscript received 18th November*, 1989)

Br. J. educ. Psychol., 60, 312-321, 1990

ATTITUDES, SUBJECTIVE NORM, PERCEIVED PREVENTIVE FACTORS, INTENTIONS AND LEARNING SCIENCE: TESTING A MODIFIED THEORY OF REASONED ACTION

By BRAHM NORWICH and JEAN DUNCAN
(Department of Educational Psychology and Special Educational Need, London University, Institute of Education)

Summary. The aims of this study were to investigate (1) a repeated measures procedure for assessing behaviour intention and learning behaviour in the field of secondary school science learning, and (2) the relationship of perceived preventive factors to other affective variables in a modified version of the Ajzen-Fishbein theory of reasoned action. Eighty-two boys and girls, aged 12-15, in a large inner city comprehensive school were assessed over seven successive science lessons in terms of: attitude, subjective norm, perceived preventive factors, overall prior learning behaviour, behaviour intention and past lesson learning behaviour. It was found that there was some correspondence between repeated classroom observation data and self and teacher reports of learning behaviour. Attitude and past learning behaviour, but not subjective norm, had a significant correlation with behaviour intention, as behaviour intention did with subsequent learning behaviour. Perceived preventive factors, however, had the strongest correlation with subsequent behaviour intention and learning behaviour and this persisted even when the contribution of other variables was partialled out. The findings are discussed in terms of the classroom assessment of these affective variables and possible modifications to a theory of reasoned action.

INTRODUCTION

Learner attitudes, beliefs and interests are often thought to influence school achievement. The study of these relationships has been popular in the areas of science and mathematics education. Reviews of such studies in the field of science attitudes and achievement have, however, found low correlations (Schibeci, 1984; Gardner, 1985). Schibeci, for example, in a meta analysis of 123 studies found a mean correlation of 0.11 between science attitudes and achievement. Much of the UK research has been concerned with general gender differences and age changes in science attitudes. From these studies it has been established that older learners have less positive attitudes to science, that boys are more interested in the physical sciences than girls, but the reverse is so for the biological sciences (Hadden and Johnstone, 1982; Kelly, 1986). US research has focused attention on factors which might influence science attitudes, such as learner, teacher and classroom environment factors. Lynn Talton and Simpson (1987), for example, found a moderate predictive relationship between attitudes to science and attitudes to classroom environment. Haladyna *et al.* (1983) found that attitudes were significantly predicted by the variables; sense of importance of science, student fatalism, teacher quality and learning environment.

The low correlations between science attitudes and achievement have been attributed partly to methodological difficulties. Schibeci (1984) notes that less care has been taken in assessing attitudes than cognitive variables, and that the relationships might be complex and subtle. This calls for the use of a theoretical framework about the nature of attitudes and how they relate to school learning and achievement. The aim of this study was to apply a particular analysis of attitudes to the field of science attitudes and achievement. In a previous study on mathematics attitudes and achievement (Norwich and Jaeger, 1989), it was assumed that engaging in certain behaviours promotes learning progress and achievement. It was proposed that these learning behaviours mediate between learners' beliefs and feelings about school subjects and achievement in those subjects.

These ideas were derived from Ajzen and Fishbein's (1980) theory of reasoned action. According to this theory, the immediate determinant of behaviour is the prior intention to perform that specific behaviour. Behavioural intention is itself influenced by two prior factors, a prescriptive and an evaluative one. The prescriptive factor, subjective norm, is the person's perception of the social pressure by important others to perform the behaviour or not. The evaluative factor, called attitudes, is the person's positive or negative affect towards the behaviour in question. The theory of reasoned action assumes that, although the importance of attitudes and subjective norm varies depending on the behaviour involved, the two factors jointly determine behaviour intention.

The theory further analyses attitudes and subjective norm in terms of underlying beliefs and values. Attitude is considered to be a function of a person's beliefs about the outcomes of the behaviour and his or her evaluations of the outcomes. Attitudes are therefore estimated by weighting each outcome belief by its corresponding evaluation and then summing these products for all salient outcome beliefs. Subjective norm is considered to be determined by whether the person believes that important others approve or not of the behaviour in question, normative beliefs, and the motivation to comply with these perceived prescriptions. Subjective norm is similarly estimated by weighting each normative belief by the corresponding motivation to comply and then summing the weighted beliefs across all the significant others.

There has been some evidence consistent with the theory in diverse fields, such as, consumer behaviour, voting, health related behaviours and smoking (Ajzen and Fishbein, 1980, Manstead *et al.*, 1984; Eiser *et al.*, 1989). Other studies have shown that intention does not always mediate the relationship between attitude and behaviour (Fredericks and Dossett, 1983). In the previous study of learning behaviours in a school context (Norwich and Jaeger, 1989), it was found that while an expectancy-value measure of attitude to learning mathematics was related to intention to engage in mathematics learning behaviours, a measure of subjective norm was not. Also, contrary to the theory, intention did not predict subsequent learning behaviour whether measured by learner self report or teacher report, once prior learning behaviour was taken into account.

The results of this initial study could be interpreted in terms of the poor validity of the measures of behaviour intention and learning behaviour, and the time period between the assessment of these factors. This was because behaviour intention was assessed nine months prior to the subsequent learning behaviour and there was a low correlation between self and teacher reported measures of learning behaviours. So, it was decided in this study to use the Ajzen-Fishbein framework with two major alterations to the assessment procedures. First, learning behaviour would be assessed by self report, teacher report and some direct classroom observation. This would enable some examination of the relationships between these three sources of evidence. Secondly, it was decided to assess the intention-behaviour link over the short period of one lesson. This would enable a sampling of repeated measures of intention and learning behaviour over several occasions.

It was also decided to elaborate the Ajzen-Fishbein model in the light of the previous finding that perceived pressures from others to engage in school learning, subjective norm, did not relate to learning intentions or behaviour. The subjective norm variable is a composite of perceived pressures from different significant others, including parents, teacher and peers. It is possible that perceived peer pressures operate most strongly in the context of classroom learning, but if this were so this influence would not be represented by the composite variable. It is also possible that teenagers perceive pressures and conditions in the classroom situation which inhibit them carrying out their intentions to learn.

As there is no reference in the literature to this alternative approach to perceived pressures and conditions, it was decided to introduce and test a new variable, termed perceived preventive factors, which might contribute to the understanding of learning intentions and behaviour. Perceived preventive factors refer to those specific obstacles which a person perceives as making it difficult to carry out the behaviour in question. In the context of school learning these obstacles could include, for example, perceived lack of capability or confidence, poor teaching and resources for learning, competing interests and/or negative influences of others.

In introducing this addition to the Ajzen-Fishbein model it is uncertain whether this variable influences intentions directly or whether it moderates the relationship between intention and behaviour. The aim of this study was therefore to explore this elaboration of the Ajzen-Fishbein approach with new assessment procedures in the field of secondary school science learning.

METHOD

Participants

The participants were 82 boys and girls from a large comprehensive school in inner London. There were four teaching groups from three different years: 20 second years aged 12/13 (9 boys and 11 girls), 15 third years aged 13/14 (all boys), 20 fourth years aged 14/15 (all boys) and 27 fourth years aged 14/15 (17 boys and 10 girls). The two fourth year groups were preparing for General Certificate of Secondary Education (GCSE) integrated science, though one group was selected to have higher science attainments. The second and third year groups were mixed ability and also followed an integrated science course. The school was located in a socially disadvantaged area and drew its pupils from a multi-ethnic population.

Design and procedure

The research was carried out over a period of seven consecutive weeks of science lessons. In the first week attitudinal and normative variables were assessed at the start of the lessons. Lessons were 70 minutes each.

These measures were repeated to assess their retest reliability in the second week. In the third week pupils completed the past learning behaviour and perceived preventive factors questionnaires. This was followed by the behaviour intention questionnaire. At the end of this lesson pupils completed the learning behaviour questionnaire. In the fourth week pupils repeated the preventive factors questionnaire and then completed the intention and behaviour questionnaires as in the previous week.

In the fifth, sixth and seventh weeks the same intention and behaviour questionnaires were completed at the start and end of the lessons. In weeks 4 to 7 two observers made observations of four selected pupils in each of the four groups. After the seven-week assessment sequence the science teachers completed the teacher overall learning behaviour questionnaire.

Pupils were informed about the purpose of the study and given the opportunity to ask questions before and while completing the questionnaires. Pupils with English as a second language were offered language help by support teachers and/or by the two researchers. When there was any indication that a pupil did not understand a questionnaire item the data were treated as missing.

Measures

To design the items for the questionnaires, a sample of 12 pupils from years 2 to 4 were interviewed about learning science, following the Ajzen-Fishbein method. The interview was designed to elicit the following information: (1) their salient beliefs about the outcomes of learning science; (2) significant others with respect to learning science; (3) what would make it difficult to learn science during lessons; (4) the kinds of behaviour in class that in their opinion would promote or impede learning science. Several science teachers were also asked about the kinds of behaviours referred to in (4).

Following the elicitation phase the questionnaires were designed on the basis of the interview responses to include the following variables:

(1) a *direct* measure of *attitude to learning science* in school. Pupils evaluated the statement "learning science in school" on three five-point bipolar scales (good-bad, pleasant-unpleasant, useful-useless). The mean evaluative rating was used as the measure;

(2) an *indirect* measure of *attitude to learning science* was based on nine outcome statements and evaluations of these outcomes, e.g., "learning science would result in my getting qualifications and a good job", "would result in my having confidence in myself". Beliefs were measured on a five-point agree-disagree scale and evaluative ratings were also assessed on a five-point good-bad scale. To obtain the belief based measure of attitude, the belief rating with respect to the outcome was multiplied by its corresponding evaluation and these products were summed;

(3) *subjective norm* was assessed by pupils indicating the extent of their agreement or disagreement with a set of normative beliefs concerning the expectations of important others, such as parents, friends, e.g., "my parents think that I should learn science". They then indicated their motivation to comply with each of these significant others, e.g.," I usually do what my parents think I should," on a five-point agree-disagree scale. Each normative belief was multiplied by the corresponding rating for motivation to comply and the sum of these products served as the overall measure of subjective norm;

(4) *perceived preventive factors* were assessed by pupils indicating on a set of 11 statements of factors which could make it difficult to learn science during science lessons. These statements were derived from the initial interview and could be grouped into (a) internal factors to do with personal capabilities, beliefs, behaviours and interests (e.g., "I will not understand what will be taught", "I will mess about in class") and (b) external factors to do with what others do (e.g., "others will distract me from learning", "the teacher will not teach well").

(5) *behaviour intention* was assessed by pupils indicating on a five-point scale whether they intended to engage in certain learning behaviours. These statements were selected for having high item-total correlations from a larger set of learning behaviours which teachers and pupils had identified (e.g., "I plan to listen carefully to the teacher," "I plan to mess about in class").

(6) *learning behaviour.* (a) questionnaire: the same six behaviour statements used for behaviour intention were used for assessing learning behaviour. There were three versions: (i) for teachers of overall learning behaviour over the last few weeks, (ii) for pupils of learning behaviour over the last few weeks, and (iii) for pupils of learning behaviour in the last lesson. (b) observation: a classroom observation schedule was designed to assess the frequency of engagement in learning behaviours. The observation categories were derived from the six behaviour statements used in the questionnaire assessment of learning behaviour. These behaviours, such as, listening, paying attention to teacher and so on, were used as indicators of engaging in learning behaviours. Observations were conducted by a main observer and a co-observer for between 15 to 20 minutes of a lesson. In each class four pupils were observed, two whom teachers identified as usually engaging in learning behaviours and two who did so less. The pupils were observed in rotation in terms of these learning behaviour indicators. Each pupil was observed for 15-second periods for the presence or absence of the indicators. When there was some doubt, the observation was recorded as uncertain. The frequencies were converted into proportions of the total time that pupils were engaged in learning behaviours.

RESULTS

Reliability measures

Cronbach alpha measures of internal and retest reliability were calculated for the variables under consideration. As Table 1 shows, the measures were at a satisfactory level for all the variables. Retest reliability correlations for the three composite variables were also at a satisfactory level. For the five lesson-based measures of intention and self reported behaviour, stability was more variable. This degree of change could be expected for measures which unlike the others are lesson specific.

TABLE 1

RELIABILITY MEASURES FOR DIFFERENT VARIABLES

		Cronbach alpha		Retest	
Composite attitude	time 1	0.79	(N=57)		
	time 2	0.76	(N=57)	0.69	(N=54)
Composite subjective norm	time 1	0.67	(N=54		
	time 2	0.79	(N=57)	0.78	(N=54)
Composite perceived preventive factors	time 1	0.89	(N=70)	0.80	(N=31)
	time 2	0.91	(N=57)		
Learning behaviour intentions	lesson 1	0.74	(N=72)		
	lesson 2	0.85	(N=65)	0.43	(N=62)
	lesson 3	0.90	(N=53)	0.74	(N=56)
	lesson 4	0.92	(N=64)	0.78	(N=48)
	lesson 5	0.83	(N=69)	0.67	(N=53)
Learning behaviour:					
i. pupils overall report		0.73	(N=70)		
ii. teacher overall report		0.91	(N=73)		
iii. pupil report	lesson 1	0.78	(N=69)		
	lesson 2	0.80	(N=63)	0.56	(N=55)
	lesson 3	0.86	(N=63)	0.52	(N=52)
	lesson 4	0.75	(N=58)	0.73	(N=47)
	lesson 5	0.82	(N=60)	0.53	(N=49)
iv observer				86–96% 92% mean	

In the subsequent analyses it was decided to use the indirect belief-based measure of attitude to learning rather than the direct measure. The direct measure had a skewed distribution with less variance and it correlated at the 0.56 level with the indirect measure.

TABLE 2

CORRELATIONS BETWEEN DIFFERENT MEASURES OF LEARNING BEHAVIOUR

	2		3	
1. Mean observation measure	0.56	(N=10)*	0.72	(N=11)*
2. Mean lesson self report	x		0.49	(N=10)**
3. Overall teacher report	x	x		
Mean lesson self report – overall teacher report r=0.53				

* $P<0.05$; ** $P<0.08$

As Table 2 shows, the average of four lesson observations per pupil correlated at a moderate but statistically significant level with the pupil and teacher self report measures. However, the correlation between the pupil and teacher self report measures was not statistically significant for the subgroup observed, but was so for the larger group.

Relationship of background variables to behaviour intention and learning behaviour

Behaviour intention and learning behaviour measures for each lesson were averaged to produce behaviour intention and learning behaviour measures. As the coefficients in Table 3 indicate, attitude and prior learning behaviour have significant low to medium correlations with subsequent behaviour intentions and learning behaviours. Subjective norm, by contract, showed no significant correlation. Average behaviour intention also correlated significantly with the subsequent average behaviour variable.

TABLE 3

Correlations Between Prior Learning Behaviour, Attitude and Subjective Norm with Mean Behaviour Intention and Learning Behaviour

	mean behaviour intention		mean learning behaviour	
Attitude	0.38*	(N=41)	0.39*	(N=39)
Subjective norm	0.19	(N=42) NS	0.10 NS	(N=38)
Prior learning behaviour	0.43*	(N=49)	0.47*	(N=44)
Mean behaviour intention – mean learning behaviour			0.65*	(N=43)

*P<0.01

Perceived preventive factors and its relationship to the Ajzen-Fishbein variables

As preventive factors was a new variable it was decided to analyse the frequency of responses to its component items.

Table 4 shows that the items with the highest levels of agreement were concerned with personal feelings, interests and the influence of friends. Capability factors concerned with skill and understanding were identified much less often.

TABLE 4

Percentage of Pupils Agreeing or Agreeing Strongly with the Following Preceived Preventive Factors (N=57)

		%	Rank
1.	Will not understand what will be taught	6.7	7=
2.	Will mess about in class	8.2	6
3.	Will not have confidence to learn	6.7	7=
4.	Will have more interest in other subjects	22.9	2
5.	Will not be able to work hard	6.5	9
6.	Will not have the skill to learn	3.2	11
7.	Others will distract me from learning	16.7	4
8.	Teacher will teach well	4.8	10
9.	Friends will encourage me to do something else	19.7	3
10.	Will want to do something else, like talk with others	11.4	5
11.	Will want to relax and take it easy	25.8	1

Two possible relationships between the perceived preventive factors and behaviour intention were investigated. The first was the possibility that preventive factors operated as a moderator variable between intention and behaviour. If this were so, one would expect that for those with high perceived preventive factors there would be no correlation between intention and behaviour. By contrast, for those with low preventive factors, an intention-behaviour correlation would be expected. The results in Table 5 show that there is no difference in the intention-behaviour correlation for those high and low in preventive factors. The other possibility is that perceived preventive factors operates on behaviour intention directly, such that those with high preventive factors will have low behaviour intention, i.e., a negative correlation. The correlational evidence is consistent with this hypothesis, showing significant medium to high correlation between preventive factors and behaviour intention and learning behaviour.

TABLE 5

CORRELATIONS BETWEEN BEHAVIOUR INTENTION AND SELF REPORTED BEHAVIOUR WHEN PERCEIVED PREVENTIVE FACTORS ARE HIGH AND LOW

	Preventive factors	
	low (uncertain or disagree)	high (agree or agree strongly)
Time 1	0.47 (N=44)*	0.42 (N=8) NS
Time 2	0.65 (N=41)*	0.73 (N=11)*

CORRELATIONS BETWEEN PERCEIVED PREVENTIVE FACTORS AND AZJEN FISHBEIN VARIABLES

	perceived preventitive factors
Mean behaviour intention	–0.71 (N=36) *
Mean learning behaviour	–0.77 (N=32) *
Attitude	–0.21 (N=52) NS
Subjective norm	–0.07 (N=52) NS
Overall learning behaviour	–0.33 (N=57) *

*P<0.01

Partial correlational analysis

Pursuing the latter finding it was decided to conduct a partial correlational analysis of behaviour intention with each of attitude, subjective norm, prior learning behaviour and perceived preventive factors while controlling for the other variables.

TABLE 6

PARTIAL AND ZERO ORDER CORRELATIONS BETWEEN ATTITUDE, SUBJECTIVE NORM, PRIOR OVERALL LEARNING BEHAVIOUR, PERCEIVED PREVENTIVE FACTORS AND BEHAVIOUR INTENTION

	with behaviour intention	
	controlling for 3 other variables	zero order
Attitude	0.10 (N=34) NS	0.38 (N=41) **
Subjective norm	0.17 (N=34) NS	0.19 (N=42) NS
Prior overall learning behaviour	0.33 (N=34) *	0.43 (N=49) **
Perceived preventive factors	–0.65 (N=34) **	–0.71 (N=36) **

*P<0.05, **P<0.01

Comparing these partial correlation coefficients with the zero order ones in Table 6 indicates that once account is taken of preventive factors and prior learning behaviour, the relationship between attitude and intention ceases to be significant. By contrast, the relationship between prior learning behaviour and intention remains significant when the relationships with the other variables are partialled out. Similarly, preventive factors still has a significant medium correlation with intention after the other variables are partialled out.

DISCUSSION

This study attempted to apply the Ajzen-Fishbein framework to the field of science learning in a follow up of some previous research on mathematics learning (Norwich and Jaeger, 1989). As in the previous study, the focus was on investigating how affective factors might relate to classroom learning behaviours. The results have some bearing on both the theoretical framework and our understanding of some of the processes of learning science in school. Conclusions do, however, have to take account of several methodological limitations.

The self report measures proved to have satisfactory internal and retest reliability. The data were, however, collected in one school with particular characteristics which make it difficult to generalise to other schools with different approaches to teaching science and with different intakes of pupils. There was also a loss of data over the period of the study which reduced the data base for statistical analysis. This loss arose for several reasons. The design involved data collection on a repeated lesson by lesson basis over seven lessons. There were some pupil absences from week to week and some pupils had difficulty in completing a form. Missing data for just one session resulted in all data for a pupil being treated as absent for computing the composite variables.

A positive assessment feature was the moderate convergence between the different sources of data on learning behaviour. The use of direct classroom observation provided reliable data on a subgroup of pupils and showed a moderate level of correspondence with teacher and pupil reports about learning behaviour. This aspect of the study was an improvement on the previous study when there was much lower correspondence between self and teacher reports. This difference could be due to the new approach to assessing learning behaviour and behaviour intention. Instead of assessing these in relation to a long time span covering several weeks of lessons, they were assessed on a lesson by lesson basis. It could be that by having a more specific assessment focus, and then averaging repeated measures over several lessons, differences are less likely to arise in the different perspectives on classroom learning behaviours.

As in the previous study, there was evidence that pupils with more positive attitudes to learning science, as measured by the belief-value procedure, tended to have more intention to engage in future learning behaviours. However, as before, subjective norm did not have a significant relationship with behaviour intention, but had one with attitude ($r = 0.42$). It was also found that those pupils who reported that they had engaged in learning behaviours in the past also had higher behaviour intentions.

The most interesting finding in this study concerns the additional variable of perceived preventive factors. This variable assumes an awareness of aspects of oneself and one's situation which can affect behaviour. In the context of the school learning of science it was assumed that pupils can identify factors which can make it difficult for them to learn — perceived preventive factors. The evidence showed that it was possible to design a measure of preventive factors which was internally reliable and stable over a short period.

Differences between pupils on this variable were mainly concerned with alternative personal interests and the adverse influence of friends and peers. Up to one in five of the pupils reported that wanting to relax/take it easy, that having more interest in other subjects and that encouragement from friends to do something else would make it hard to learn science. It is not unexpected, then, to find that those high in preventive factors had low behaviour intention and learning behaviour. Preventive factors also had a negative but low relationship with attitude to learning and overall past learning behaviour, though only the latter was statistically significant. There was no evidence that preventive factors moderated the influence of intention on behaviour, though this could be due to the use of a low number of pupils in the analysis with high preventive factors.

The lack of relationship between subjective norm and perceived preventive factors is notable because both contain items related to the perceived influence of peers on science learning, but only preventive factors were related to intention. This could be understood in terms of subjective norm being a composite variable of perceived pressure from significant others. As this measure included perceived adult pressure, the contribution from perceived pressure from peers may have been lessened. Perceived preventive factors similarly contained items which referred to other factors affecting science learning. Secondly, the subjective norm variable, which involves perceived pressures to learn science, is different from the focus of the perceived preventive factors variable which involves factors making it hard to learn science. These findings are therefore relevant to how perceived pressures are conceptualised and assessed in the Ajzen-Fishbein theory.

In the partial correlational analysis, it was found that by controlling for preventive factors and prior learning behaviour there was no longer a significant relationship between attitude to learning science and intention to learn. This analysis did, however, show that past learning behaviour and perceived preventive factors continue to have significant independent relationships to behaviour intention.

This study has significance for understanding attitude to school learning, particularly in the area of learning science. It illustrates how specific beliefs, values and intentions relate to actual classroom behaviour. The role of perceived preventive factors appeared to be central to understanding differences between pupils in how they set behaviour intentions and their subsequent classroom behaviour. The study also shows the significance of beliefs and values about the consequences of learning science to learning intentions. As regards the Ajzen-Fishbein theory of reasoned action, it illustrates three significant points. Firstly, there was further evidence that subjective norms were not independent of attitudinal factors and that behaviour intention was best understood in terms of attitude alone. Secondly, the more specific time-focused procedure of assessing intention and behaviour resulted in more reliable measures. This proved to be an improvement on the broader time focus used in the previous study (Norwich and Jaeger, 1989). Thirdly, the introduction of the variable, perceived preventive factors, illustrated how the theory of reasoned action could be elaborated to take account of factors which might influence behaviours of a repeated and skill-based nature in an institutional context. This suggests that the Ajzen-Fishbein framework may need to be adapted to include the influence of other personal interests and immediate social pressures on classroom intentions and behaviours.

Correspondence and requests for reprints should be addressed to Dr Brahm Norwich, EPSEN, London University, 25 Woburn Square, London WC1H 0AA.

REFERENCES

Ajzen, I., and Fishbein, M. (1980). *Understanding Attitudes and Predicting Social Attitudes*. Englewood Cliffs, N.J.: Prentice Hall.

Eiser, J. R., Morgan, M., Gammage, P., and Gray, E. (1989). Adolescent smoking, attitudes, norms and parental influence. *Br. J. soc. Psychol* , 28, 3, 193-202.

Fredericks, A. J., and Dossett, D. L. (1983) Attitudes-behaviour relations: a comparison of the Fishbein-Azjen and Bentler-Speckart models. *J. Pers. soc. Psychol.*, 45, 501-512.

GARDNER, P. L. (1985). Student interests in science and technology: an international review. In LEHRKE, M., HOFFMAN, L., and GARDNER, P. L. (Eds.), *Interest in Science and Technology*. Kiel: IPN.

HADDEN, R. A., and JOHNSTONE, A. H. (1983). Secondary school pupils' attitudes to science: the year of decision. *Eur. J. Sci. Educ.*, 54, 429-438.

HALADYNA, T., OLSEN, R., and SHAUGHNESSY, J. (1983). Correlates of class attitude to science. *J. Res. Sci. Teaching*, 20 (4), 311-324.

LYN TALTON, E., and SIMPSON, R. D. (1987). Relationships of attitudes to classroom environment with attitude to and achievement in science among tenth grade biology students. *J. Res. Sci. Teaching*, 24 (6), 507-516.

KELLY, A. (1986). The development of girls' and boys' attitudes to science: a longitudinal study. *Eur. J. Sci. Educ.*, 8 (4), 399-412.

MANSTEAD, A. S. R., PLEVIN, C. E., and SMART, J. L. (1984). Predicting mothers' choice of infant feeding method. *Br. J. soc. Psychol.*, 23, 223-231.

NORWICH, B. and JAEGER, M. (1989). The predictive relationship between beliefs, attitudes, intentions and secondary school mathematics, learning: a theory of reasoned action approach. *Br. J. educ. Psychol* 59, 3, 314-325.

SCHIBECI, R. A. (1984). Attitudes in science; an update. *Stud. Sci. Educ.*, 11, 26-59.

(*Manuscript received 2nd February*, 1990)

Br. J. educ. Psychol., 60, 322-329, 1990

STUDENT PERCEPTIONS OF PERMISSIVENESS AND TEACHER-INSTIGATED DISCIPLINARY STRATEGIES

By NEVILLE J. KING
(Monash University)
ELEONORA GULLONE
(Victoria College)
AND MARK R. DADDS
(University of Queensland)

SUMMARY. Using a self-report questionnaire, we examined the perceptions of 616 primary and secondary school students of permissiveness and disciplinary strategies. Acceptability ratings formed the dependent variable. The most acceptable strategy was individual student-teacher discussion, followed by exclusionary time out, then physical punishment, and least acceptable was permissiveness. Also examined was the influence of the misbehaviour situation and evaluator characteristics (age and gender). Several methodological limitations are discussed including the range of disciplinary strategies that were assessed, and focus on an individual rather than group of students.

INTRODUCTION

SINCE Wolf's (1978) now classic article on social validation, the acceptability of intervention strategies has been of growing concern to behaviour therapists. There are several reasons underlying the current interest in the acceptability of behavioural interventions (Gullone and King, 1989). Lebow (1982) maintains that research on consumer satisfaction is part of the trend towards a more consumer-oriented society. There also has been an increase in concern over the rights of individuals who are treated. Related to this issue are informed consent, and ethical and legal considerations (Braun, 1975; Budd and Baer, 1976; Kazdin, 1977). Crowe *et al.* (1972), amongst others, maintain that an acceptable intervention programme is more likely to be implemented and adhered to than one which is less acceptable, although the latter may be potentially more effective. Wolf (1978) is in agreement with Crowe *et al.* (1972) when he states that "it may be that not only is it important to determine the acceptability of treatment procedures to participants for ethical reasons, it may also be that acceptability of the programme is related to effectiveness, as well as to the likelihood that the programme will be adopted and supported by others" (p. 210).

Children's perceptions are of particular interest as youngsters are often not given the opportunity to express opinions about the acceptability of interventions (Gullone and King, 1989). Recent studies by Dadds and his colleagues evaluated the acceptability of five maternal disciplinary techniques including permissiveness, physical punishment, directed discussion, quiet time, and time-out (Dadds *et al.*, 1987). These techniques were evaluated across four different situations, i.e., non-compliance with an initiating instruction, aggression toward others, non-compliance with a terminating instruction, and non-compliance with a known rule. Twenty children, aged 4 to 8 years, participated in this investigation. They were classified as either behaviour problem or non-problem children. Comparisons were also made between the behaviour problem children's perceptions of technique acceptability before and after behavioural treatment. Using a three-point "face rating scale", ranging from very wrong to very right, Dadds *et al.* found that the children primarily discriminated between intervention and non-intervention rather than between the different techniques. Out of the five techniques only permissiveness was rated as unacceptable. A further finding was that exposure to behavioural disciplinary techniques (e.g., differential reinforcement, exclusionary and non-exclusionary time-out) did not alter children's acceptability ratings of the five disciplinary techniques. In addition to children's perceptions of maternal disciplinary strategies, we also must examine the perceptions of children in other settings.

School settings are of particular relevance, especially in view of the importance of discipline to social and academic development. There is much controversy about discipline in our schools and about what teachers and others should be doing to foster discipline. This debate has been fuelled by the restriction or elimination of corporal punishment in schools. Consequently, many believe that the behaviour of children and adolescents has emerged as the most serious educational problem in our schools. Perhaps surprisingly, a dearth of research has addressed the perceptions of students themselves about disciplinary matters. From the few studies that have been reported (e.g., Elliot *et al.*, 1986; Turco and Elliot, 1986), it appears that students differentiate between disciplinary procedures. However, whether students actually prefer discipline to permissiveness has not been examined. Intuitively, we might predict that students would prefer their teachers to be permissive and not set limits on their behaviour. On the other hand, the findings of Dadds *et al.* (1987) in relation to preschool children's perceptions of maternal disciplinary strategies suggest that students may actually prefer their teachers to be firm in classroom behaviour management. Therefore, the major aim of the present investigation was to examine the perceptions of students towards permissiveness and teacher-instigated disciplinary strategies.

Consistent with the behavioural tradition in the study of personality and social learning (Mischel, 1971), it might be argued that students' attitudes or perceptions on these matters are situation-specific. Thus, acceptability evaluations may vary depending on the type of misbehaviour in question (e.g., non-compliance to a teacher request versus disruptive behaviour in class). Although the data are fairly sparse, it also has been suggested that developmental factors may influence acceptability (Elliot, 1987). Hence, for example, younger children might be expected to prefer more direct intervention with older children preferring greater autonomy. Further, the rater's gender may play a role in acceptability evaluations. In our culture, males and females are expected to develop different attitudes and behaviours in accord with sex-role stereotypes (Sarason *et al.*, 1960). Consequently, the second aim of the study was to assess the extent to which students' perceptions towards disciplinary strategies are influenced by situational factors and evaluator characteristics (age and gender).

METHOD

Sample

The sample consisted of 616 primary and secondary school students, aged 8 to 16 years. There were 209 participants aged between 8 and 10 years (104 boys and 105 girls), 192 aged between 11 and 13 years (108 boys and 84 girls), and 215 aged between 14 and 16 years (124 boys and 91 girls). The schools were from lower to middle socio-economic areas in Melbourne, Victoria.

Instrument

We developed a self-report questionnaire (Discipline Questionnaire) to assess students' perceptions towards permissiveness and disciplinary strategies. Similar to the interview instrument used by Siegal and Cowen (1984) and Dadds *et al.* (1987), the Discipline

Questionnaire requires the child to rate the behaviour of a hypothetical adult (teacher) across different situations in response to a student misbehaving. The situations presented to the respondents included: (1) non-compliance with an instruction to pick up books, (2) aggression towards another student, and (3) temper tantrum in the classroom. Each of these situations was illustrated on the questionnaire.

The disciplinary techniques the teachers were described as using entailed: (1) permissiveness — the teacher "let the child get away with it", (2) physical punishment — the teacher "smacked the child", (3) directed discussion — the teacher "talked to the child", and (4) time-out (exclusionary) — the teacher "put the child in another room"). Students were required to circle whether the teacher was *Very Wrong, Half Right — Half Wrong*, or *Very Right* (scored 0, 1, 2). As a result of some initial pilot work on the questionnaire, this simple three-point rating scale was used to enhance the participation of students (especially younger students). There were evaluations of four teacher behaviours over three situations thus providing 12 items in total.

Procedure

The Discipline Questionnaire was administered to students during class. A research assistant explained that "We want to know what students think about discipline and what should be done in schools. You can help by completing this questionnaire". The primary school students were required to keep pace with the research assistant who explained each item. The secondary school students completed the questionnaire with little assistance from the research assistant. Students with reading difficulties were given individual assistance on the stimulus material and responses. Approximately 640 questionnaires were distributed to students. However, the responses were unclear on some of the questionnaires thus leaving a final sample of 616.

RESULTS

Students' perceptions of teacher disciplinary strategies were analysed using a 2 (evaluator gender) x 3 (evaluator age group) x 3 (misbehaviour situation) x 4 (teacher disciplinary strategy) multivariate analysis of variance. Both Wilks' lambda and F values are reported for the interactions and main effects (SPSS-X, 1988). Significant multivariate effects were further analysed using univariate analyses of variance. Table 1 shows the acceptability means and SDs for each of the disciplinary strategies across the three situations. The means indicate a preference for directed discussion followed, in order, by time-out, physical punishment, and permissiveness.

TABLE 1

MEANS AND SDs FOR STUDENTS' ACCEPTABILITY RATINGS

	Situation 1		Situation 2		Situation 3		Overall	
	M	SD	M	SD	M	SD	M	SD
Permissiveness	0.37	0.57	0.28	0.58	0.32	0.59	0.32	0.58
Physical Punishment	0.37	0.63	0.41	0.67	0.47	0.70	0.42	0.67
Time-out	0.78	0.64	0.91	0.67	1.01	0.72	0.90	0.68
Directed Discussion	1.72	0.53	1.55	0.64	1.64	0.59	1.64	0.59

Two-way interactions were found between age of the evaluator and strategy [Wilks' lambda = 0.90, F(6,1186) = 11.00, P<0.001], between gender of the evaluator and disciplinary strategy [Wilks' lambda = 0.95, F(3,593) = 9.52, P<0.001] as well as between situation and strategy [Wilks' lambda = 0.86, F(6,590) = 15.97, P<0.001].

A main effect was found for disciplinary strategy [Wilks' lambda = 0.21, F(3,593) = 724.20, P<0.001] and for situation [Wilks' lambda = 0.94, F(2,594) = 17.75, P<0.001] while no significant main effects were found for evaluator age or gender.

The interaction between age of the evaluator and disciplinary strategy was the result of the 11-13 year olds and 14-16 year-olds rating time-out and permissiveness to be more acceptable while physical punishment and directed discussion were rated to be less acceptable than were rated by the 8-10 year-old students (see Figure 1).

The interaction between gender of the evaluator and disciplinary strategy was due to females' acceptability evaluations being higher than those of males for directed discussion and time-out and lower for permissiveness and physical punishment, in all three situations (see Figure 2).

The interaction between misbehaviour situation and disciplinary strategy was the result of physical punishment and time-out being rated as most acceptable in situation three (temper tantrum), followed by situation two (hitting another child), and least acceptable in situation one (refusing to pick up books). In contrast, permissiveness and directed discussion were perceived to be most acceptable in situation one followed by situation three and least acceptable in situation two.

Univariate F-tests, computed for each situation separately, showed that directed discussion was evaluated to be significantly more acceptable than each of the other strategies and time-out was found to be significantly more acceptable than physical punishment and permissiveness. These univariate contrasts were found to be significant at alpha = 0.001 for each of the three situations.

DISCUSSION

As already noted, similar findings concerning permissiveness have been reported on preschool children (Siegal and Cowen, 1984; Siegal and Barclay, 1985; Dadds *et al.*, 1987). Thus, the present findings may be an expression of an attitude that is formed in early years of life as a result of child-rearing practices. Although beyond the scope of our discussion, the study of moral and social development may also be relevant to this particular finding.

However, it was not a simple matter of students preferring intervention strategies to non-intervention (permissiveness) as physical punishment and exclusionary time-out were not acceptable disciplinary strategies either. Given the aversive nature of physical punishment and time-out, these evaluations are not surprising (cf. Kazdin *et al.*, 1981). Of course, exclusionary time-out is frequently used in behavioural classroom management. Therefore, teachers and others might inform students about the rationale and procedural aspects of exclusionary time out as a means of enhancing its acceptability (cf. Singh and Katz, 1985). Further, it should be stressed that time-out encompasses procedures beyond the isolation or exclusion of the subject. It is likely that exclusionary time-out will have greater acceptance when used in conjunction with more positive behavioural procedures.

O'Leary and his colleagues have differentiated between loud and soft reprimands in the control of disruptive students (O'Leary *et al.*, 1970). In contrast to the loud reprimand, a soft reprimand entails the teacher approaching the individual student and speaking in a soft manner so that other children cannot hear the reprimand. Our finding on directed discussion might be interpreted as social validation for the soft reprimand, a technique that teachers have found to be viable and effective. Similarly, Elliot *et al.*

FIGURE 1

MEAN ACCEPTABILITY RATINGS OF DISCIPLINARY STRATEGIES BY AGE

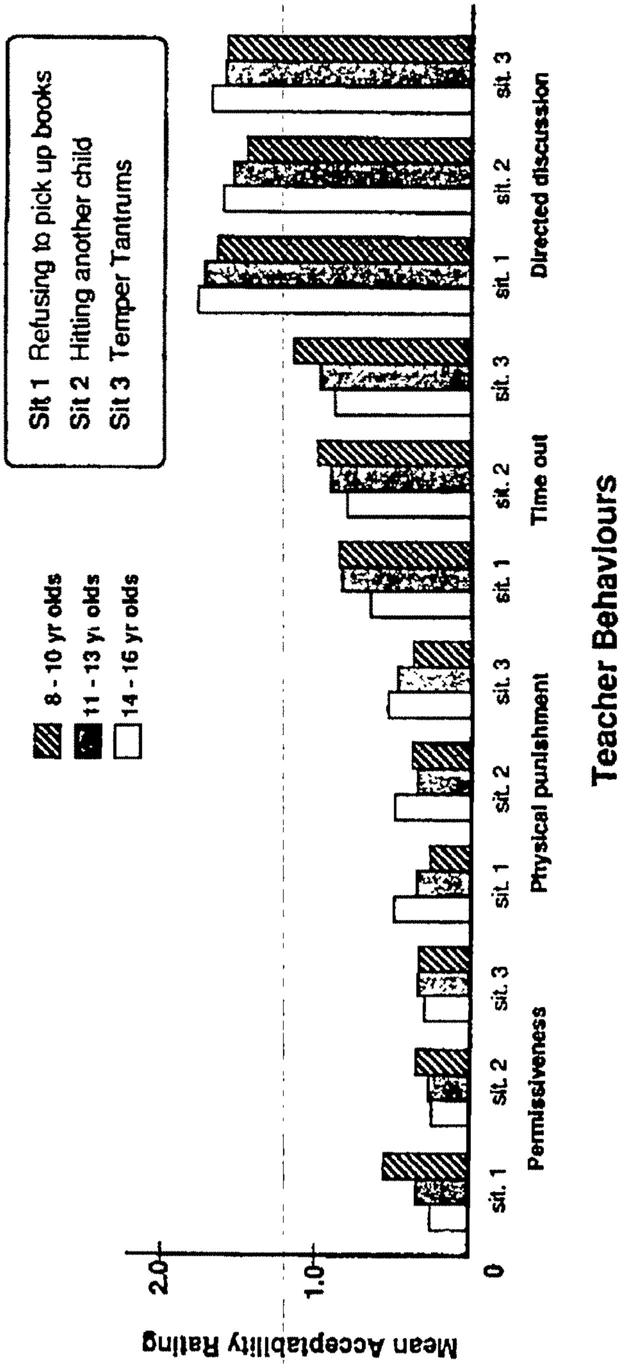
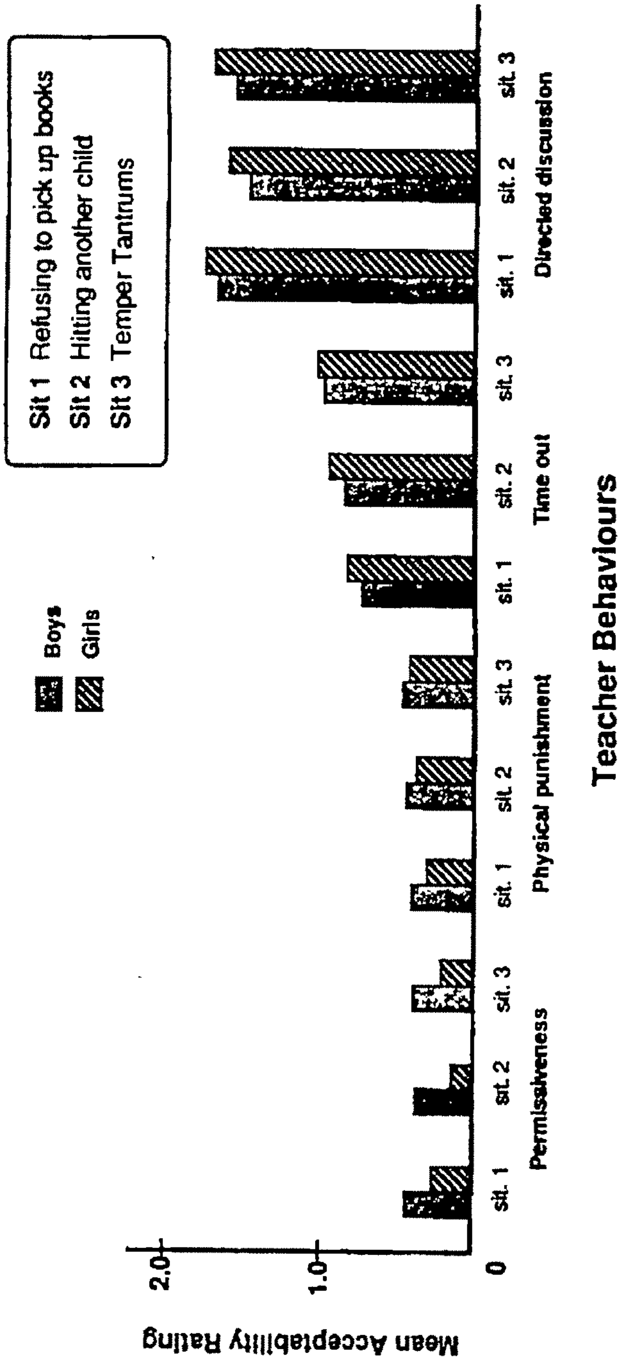


FIGURE 2

MEAN ACCEPTABILITY RATINGS OF DISCIPLINARY STRATEGIES BY GENDER



(1986) found that interventions that emphasised individual teacher-student interactions were rated as most acceptable.

In general, our findings on situational factors and evaluator characteristics are complex and require further research. Regards gender differences, we speculate that these findings may reflect subtle differences between boys and girls in terms of sex-role stereotyping. As part of being "masculine", boys are expected not only to be more aggressive than girls, but also to be subjected to harsher disciplinary treatment than girls. On the other hand, as part of being "feminine", girls are expected to be more passive and compliant than boys, and thus not require the same degree of physical discipline.

Any implications of our findings for the application of behavioural strategies in school settings are tempered by several observations. First, it will be recognised that students were presented with several disciplinary strategies for evaluation. However, far more disciplinary strategies are used by teachers (see, for example, Alberto and Troutman, 1982). Given that students differentiated between the disciplinary strategies presented to them, care should be taken in making generalisations to other teacher-instigated disciplinary strategies. Second, in the present investigation, acceptability ratings were examined in relation to an individual student. However, the classroom setting typically involves a group of students. Whether the teacher focuses on the individual student or group appears to be a fairly critical dimension in the acceptability evaluations of students (see Elliot *et al.*, 1986). Third, our study examined the perceptions of student towards interventions in relation to another student. Perhaps different findings would have emerged if respondents had been required to evaluate the disciplinary strategies in relation to themselves. Notwithstanding these limitations, our findings clearly suggest that students hold definite opinions about permissiveness and teacher-instigated disciplinary strategies.

Correspondence and requests for reprints should be addressed to Dr N. J. King, Faculty of Education, Monash University, Clayton, Victoria, Australia 3168.

REFERENCES

Alberto, P. A., and Troutman, A. C. (1982). *Applied Behavior Analysis for Teachers. Influencing Student Performance*. Columbus, Ohio: Merrill.

Braun, S. H. (1975). Ethical issues in behavior modification. *Behav. Therapy*, 6, 51-62.

Budd, K. S., and Baer, D. M. (1976). Behavior modification and the law: implications of recent judicial decisions. *J. Psychiat. Law*, 4, 171-244.

Crowe, M. J., Marks, I. M., Agras, W. S., and Leitenberg, H. (1972). Time-limited desensitisation, implosion and shaping for phobic patients: a cross over study. *Behav. Res. Therapy*, 10, 319-328.

Dadds, M. R., Adlington, F. M., and Christensen, A. P. (1987). Children's perceptions of time out and other maternal disciplinary strategies: the effects of clinic status and exposure to behavioural treatment. *Behav. Change*, 4, 3-13.

Elliot, S. N. (1987). Acceptability of behavioral treatments: review of variables that influence treatment selection. *Profess. Psychol.: Res. Pract.*, 19, 68-80.

Elliot, S. N., Witt, J. C., Galvin, G. A., and Moe, G. L. (1986). Children's involvement in intervention selection: acceptability of interventions of misbehaving peers. *Profess. Psychol.: Res. Pract.*, 17, 235-241.

Gullone, E., and King, N. J. (1989). Acceptability of behavioral interventions: child and caregiver perceptions. In Hersen, M., Eisler, R. and Miller, P. M. (Eds.), *Progress in Behavior Modification, Vol. 24*. Newbury Park, CA: Sage.

Kazdin, A. E. (1977). Assessing the clinical or applied importance of behavior change through social validation. *Behav. Modif.*, 1, 427-452.

Kazdin, A. E., French, N. H., and Sherick, R. B. (1981). Acceptability of alternative treatments for children: evaluations by impatient children, parents, and staff. *J. consult. clin. Psychol.*, 49, 900-907.

Lebow, J. (1982). Consumer satisfaction with mental health treatment. *Psychol. Bull.*, 91, 244-259.

Mischel, W. (1971). *Introduction to Personality*. New York: Holt, Rinehart and Winston.

O'Leary, K. D., Kaufman, K. F., Kass, R. E., and Drabman, R. S. (1970). The effects of loud and soft reprimands on the behavior of disruptive students. *Except. Children*, 37, 145-155.

Sarason, S. B., Davidson, K. S., Lightall, F. F., Waite, R. R., and Ruebush, B. K. (1960). *Anxiety in Elementary School Children*. New York: Wiley.

Siegal, M., and Barclay, M. S. (1985). Children's evaluations of fathers' socialisation behavior. *Dev. Psychol.*, 21, 1090-1096.

SIEGAL, M., and COWEN, J. (1984). Appraisal of intervention: the mother's versus the culprit's behavior as determinants of children's evaluations of discipline techniques. *Child Dev.*, 55, 1760-1766.
SINGH, N. N., and KATZ, R. C. (1985). On the modification of acceptability ratings for alternative child treatments. *Behav. Modif.*, 9, 375-386.
SPSS Inc. (1988). *SPSS-X User's Guide* (3rd ed.). Chicago: SPSS Inc.
TURCO, T. L., and ELLIOT, S. M. (1986). Assessment of students' acceptability ratings of teacher-initiated intervention for classroom misbehavior. *J. School Psychol.*, 24, 277-283.
VAN HOUTEN, R. (1979). Social validation: the evolution of standards of competency for target behaviors. *J. appl. Behav. Anal.*, 12, 581-591.
WOLF, M. M. (1978). Social validity: the case for subjective measurement or how applied behavior analysis is finding its heart. *J. appl. Behav. Anal.*, 11, 203-214.

(*Manuscript received 18th November*, 1989)

Br. J. educ. Psychol., 60, 330-337, 1990

TIME ON EDUCATIONAL ACTIVITIES AT HOME AND EDUCATIONAL PROGRESS IN INFANT SCHOOL

By IAN PLEWIS, ANN MOONEY AND ROSEMARY CREESER
(Thomas Coram Research Unit, Institute of Education, University of London)

SUMMARY. This paper reports on the amount of time 6-year-olds in inner London engage in educational activities at home, and links these results with earlier research on the association between parental help with reading and children's reading progress. The activities data were collected using a time budget approach. The median child in this sample read aloud at home about 20 minutes per week, and received about 40 minutes per week total "teaching" in reading, writing and maths outside school. Black (Afro-Caribbean) children read aloud more than white children, girls read on their own more than boys, and the more educational qualifications the mother had, the more time the children spent on all the measured educational activities. Associations between reading activities at home and reading attainment at the end of year 1 ranged from 0.13 to 0.27 but associations with progress were small. However, children who spent more time with books at home did make greater reading progress during year 1. Implications for understanding social class differences in educational attainment and for designing intervention studies are discussed.

INTRODUCTION

THE educational attainments of working class children are, as a group, lower than those of their middle class peers. This is a well-attested finding from educational research (see, for example, Osborn and Milbank, 1987). For many years, researchers and practitioners alike have sought ways of improving the educational performance of working class children (e.g., Halsey, 1972). Reading has been a particular focus of their concern with considerable emphasis given in the last decade to the importance of parental help as a means of raising reading levels. Considerable optimism about the efficacy of parental help, more specifically parents listening to their children reading aloud, was generated first by a survey (Hewison and Tizard, 1980) and then by the ensuing experimental study often known as the Haringey project (Tizard *et al.*, 1982). After two years of the Haringey intervention, during which parents were encouraged by their child's teacher and by the research team to listen to their child read, 66 per cent of the children were reading at or above the national average compared with 29 per cent in parallel control classes in the same schools. Only 6 per cent were poor readers (test scores below 85) compared with 17 per cent in the control classes. Hewison (1988) reported that differences were maintained three years after the end of the intervention.

There has not, however, been an accumulation of research findings which confirm the efficacy of parental help. Instead, Hannon (1987), who compared reading scores of children whose parents were encouraged to hear them read with the scores of cohorts of pre-project children who had not been so encouraged found no significant improvement (although his survey results were in line with earlier findings). Moreover, Tizard *et al.* (1988) found that parental help was not related to the reading test scores of inner London infant school children. Nor was there evidence of greater progress for those children whose parents said they regularly listened to their child read compared with those who said they did so infrequently or not at all. However, Tizard *et al.* did find a higher reported level of parental help than Hewison and Tizard had found a decade earlier.

Attempts have been made to reconcile these results. Thus, Tizard *et al.* argue that parental help is only useful if it is part of a carefully planned approach by the school and the child's teacher, and Hannon (1989) has suggested that standardised reading tests may not be the most appropriate outcome variable for assessing the effectiveness of parental help. Moreover, as Hewison (1988) points out, the success of the Haringey project might be attributed to the home-visiting by the researchers and to the nature of the population at which it was directed. It is also possible that the way in which data on parental help at home have often been collected, by asking parents directly how much time they devote to it, may have led to an over-estimation of the amount of help given. This in turn could obscure associations with attainment and progress in non-experimental studies. Finally, it must be recognised that the Haringey study was the only one employing any kind of randomisation.

This was the context for the study described here. We wanted to find out more about the accuracy of responses to the direct questions to parents which were used in earlier studies, and whether stronger associations between reading attainment and progress on the one hand, and the amount of parental help on the other, would be found with a more accurate data collection method. The study had the following aims:

(a) to obtain accurate data at the individual, as well as the aggregate, level on the extent to which year 1 pupils (aged 5 to 6) are helped with reading, writing and maths (the '3Rs') by their parents and others outside school.

(b) to relate 3R activities outside school to socio-demographic variables such as ethnic group, sex, mother's education and marital status.

(c) to relate reading and maths activities to pupils' progress in those subjects in year 1.

METHOD

Here we give a brief description of our method and sample. Further details, particularly about the reliability and validity of our approach, can be found in Plewis *et al.* (1990).

Data collection

Direct questions were not used in this study. Data were collected using a "time budget" approach. Information was collected on the sample children's activities outside school for the previous day, using the phone as much as possible. It was thought that 6-year-olds were too young to be interviewed directly in this way and so one of their parents or guardians, usually their mother, responded for them. The interviewers went through the 24 hours of the previous day or, on Mondays, the previous week-end, and asked the parent to give a sequential account, with timings, of what the child had been doing during that time. Interviewers probed to get as much detail as possible, particularly at those times of the day when educational activities might have taken place, if only for a short time.

Following a pilot study of 40 children during the autumn term of year 1, it was decided that data of sufficient reliability at the child level could be obtained by collecting data for three separate "days" (a weekend counting as a day here). Each respondent was assigned a pattern of contact days (not dates) and the three white, female interviewers were only permitted to collect data for these days, although the order was not important. As far as possible, there was an interval of two weeks between each interview. All contacts with a respondent were by the same interviewer. From the pilot study it was clear that telephone interviewing was feasible and 77 per cent of the contacts in the main study were by phone. Data were only collected during term time (mostly the spring term) when a large proportion of the child's week was spent at school. All respondents were sent an advance letter from their child's school.

Sample

The children were chosen from a sample of 20 inner London infant schools which had been selected for a parallel study. Eighteen of the schools had been in the earlier longitudinal study described in Tizard *et al.* (1988). From 19 of these schools (one school did not want to participate in this project), 230 children were selected. As in the earlier project, the sample was restricted to two ethnic groups — white children whose parents were born in the UK and black British children of Afro-Caribbean origin. An attempt was made to select at least one white boy, one white girl, one black boy and one black girl from each school and to have approximately equal numbers from these groups in the sample as a whole. It should be noted that the sample is not a probability sample of children attending London infant schools.

Table 1 shows the overall response rate. Also 81 per cent of those receiving a first contact had all three contacts.

TABLE 1

RESPONSE RATE

		Response rate (%)
Addresses issued	230	
Refusals (N=5), never contacted (N=17), child too old etc (N=12)	34	
Final sample (first contact)	196	90
Second contact	187	86
Third contact	158	72

Measures

The frequency and duration of the following 3R activities were coded: reading aloud; reading on own; being read to; writing on own; writing with other; maths on own and maths with other. It was necessary to make some essentially arbitrary decisions about coding. For example, flipping through books and comics was coded as reading on own and for those occasions when the respondent reported that the child was involved in more than one activity, e.g., watching TV and looking at a book, then a fraction of the period was allocated to the 3R activity. There were some activities, particularly those that children did on their own, when it was only possible to code that it happened and not the duration. There were occasions, too, when children were out of the home with relatives or friends when it was not possible to find out how they spent their time.

To measure attainment and progress in reading and maths, children were tested at the end of their reception year (i.e., before the activities data were collected) and at the end of year 1 (summer term, 1988). At the end of the reception year, reading and maths tests were administered to children individually. The reading test comprised measures of children's knowledge of concepts about print, their ability to identify letters and to recognise words; the maths test was the subtest of the British Ability Scales. At the end of year 1, the Young's Reading and Young's Maths tests were administered to the class. For both years, these were the same tests as those used by Tizard *et al.* (1988).

RESULTS

Time spent in different 3R activities

Table 2 shows how much time was spent each week in different 3R activities. These estimates were obtained by weighting up the daily estimates; they include the data from the pilot study and also from those children for whom data were collected on fewer than three occasions. Consequently, they are less affected by any non-response bias, but are more variable than estimates based on children with complete data. Children were read to more frequently at weekends than during the week and the weighting reflected this, but there was no variation by day of the week for the other variables. "Time spent with books" in Table 2 is the sum of the three reading variables; "total teaching" is the sum of "reading aloud", "writing with other" and "maths with other" where "other" means parent, other adult or older sibling.

TABLE 2

TIME SPENT IN 3R ACTIVITIES (MINUTES/WEEK)

	N	Median	Mean	SD	None (%)
Reading aloud	228	18.6	41.8	67.5	40
Reading on own	221	12.9	38.3	62.6	44
Read to	226	11.0	44.8	65.5	48
Time spent with books	210	93.7	122.6	126.0	16
Writing	203	0	37.6	71.4	53
Maths	229	0	14.8	32.7	69
Total teaching	223	39.6	65.8	85.0	28

The average (median) child in this sample reads aloud at home for less than 20 minutes per week, spends about one and a half hours with books each week and receives about 40 minutes "total teaching" outside school.

Associations with socio-demographic variables

The following socio-demographic variables were considered: ethnic group [white (51 per cent) and black (49 per cent)], sex, marital status [one- (41 per cent) and two-parent (59 per cent) families] and mother's education [no qualifications, (39 per cent), up to 0-level (45 per cent), A-level and above (16 per cent)]. Because the data on the amount of time spent in different activities were skewed, group differences were tested with Wilcoxon and Kruskal-Wallis non-parametric tests. Differences between groups in the frequency of an activity (i.e., days per week) were assessed using multivariate logistic models for ordered responses (McCullagh, 1980).

TABLE 3

MEDIAN TIME SPENT (MINUTES/WEEK) BY ETHNIC GROUP

	Black	White	P<
Reading aloud	26	8	0.03
Read to	0	34	0.02
Total teaching	51	30	0.05

Black children read aloud more than white children, white children were read to more than black children (but see below). There were no ethnic group differences for reading on own nor for time spent with books, writing and maths. (Reading on own included looking at a book even though the child may have been unable to read it.) Black children scored higher than white children on total teaching, mainly because this variable is highly correlated with reading aloud ($r_s = 0.81$). Table 3 gives the medians and significance levels for those variables for which the ethnic differences were statistically significant.

TABLE 4

MEDIAN TIME SPENT (MINUTES/WEEK) BY SEX

	Boys	Girls	P<
Reading on own	0	27	0.01
Time spent with books	69	112	0.02

Girls read on their own more than boys. The differences for the other two reading variables were small but girls spent more time with books than boys (Table 4). There were no sex differences for the total amount of time in writing and maths although girls engaged in these activities more frequently than boys ($P<0.01$ for writing, $P<0.04$ for maths). There was no evidence for ethnic group sex interactions for any of the 3R activities.

TABLE 5

MEDIAN TIME SPENT (MINUTES/WEEK) BY MOTHER'S EDUCATION

	>A-level	Up to O-level	None	P<
Reading aloud	40	20	0	0.001
Read to	68	11	0	0.001
Time spent with books	157	98	58	0.001
Total teaching	66	45	28	0.01

Associations with marital status were generally small except for reading to children, where children in two-parent families scored higher ($P<0.001$). This association with marital status accounts for the ethnic group difference; there was no ethnic difference for being read to within one- and two-parent families. Associations with mother's education were, however, substantial, particularly for reading aloud, being read to (and hence time spent with books) and total teaching (see Table 5). The more educational qualifications a mother had, the higher the child's score on those variables.

Associations with attainment and progress

Here we look at associations between the reading test at the end of year 1 and the four reading variables, and between the sum of the reading and maths test scores and the total teaching measure. As time spent in maths activities was unreliably measured at the individual level (see Plewis *et al.*, 1990), and as writing was not tested, associations with writing and maths activities were not considered. The Spearman correlations between the year 1 reading test and the weekly estimates of reading aloud, being read to, reading on own and time spent with books were 0.27, 0.13, 0.19 and 0.27 respectively. The correlation between the total test score and total teaching was 0.22. Thus, the more children engaged in 3R activities outside school during year 1, the higher their test scores at the end of year 1.

However, it is more useful to look longitudinally at the relation between measures of 3R activities and progress during year 1, rather than at cross-sectional associations with attainment. The appropriate statistical methods for these analyses, incorporating corrections for measurement error, are set out in the Appendix. Essentially, we looked at associations between the year 1 test scores and activities, controlling for the test score at the end of reception. All the results were in the expected direction — higher amounts of the activity are associated with greater progress. However, the sizes of the effects for reading aloud and being read to, and for total teaching, were small and neither educationally nor statistically significant. They were somewhat larger (and significant at the 0.05 level) for the amount of reading on own (although not for frequency), and for time spent with books. The difference in progress between those children at the upper quartile and those scoring zero for the activity was about 1/4 SD of the year 1 reading test for reading on own (where the upper quartile was 50 mins/week) and about 1/2 SD for time spent with books (where the upper quartile was 176 mins/week).

DISCUSSION

We set out in this study to relate 3R activities at home to socio-demographic variables, and to attainment and progress. Our finding that black parents hear their children read more than white parents is consistent with results in Tizard *et al.* (1988) who found that parents of black children attending multi-ethnic inner London schools did more teaching of reading than parents of white children attending the same schools, as well as being more positive about helping at home and having higher aspirations for their children. We found girls spent more time with books than boys. This is also consistent with a finding of Tizard *et al.* even though their variable — experience with books — was not directly comparable with our own. As we found no evidence of any ethnic group sex interactions for any of the reading activities at home, we are not able to throw any light on why black boys tend to fall behind in reading over the infant school period (documented in Tizard *et al.*, 1988).

However, our strongest associations were between measures of 3R activities and mother's educational qualifications. If we use educational qualifications as a proxy for social class, we can say middle class children did a lot more 3R activities at home. Maybe social class differences in attainment can be explained by the fact that middle class parents devote more time to educational activities at home. However, after controlling for time spent with books, reading attainment at the end of the year was still associated with mother's level of education. So, although not accounting for all of the difference, more time in these activities does explain some of it.

Why do middle class parents give more time to 3R activities? There are several possible explanations. One which has been favoured in the past is that middle class parents have more favourable attitudes to education than working class parents and consequently spend more time with their children in 3R activities (CACE, 1967). This may of course be true, but one must question earlier research methodology in this area. It was teachers rather than parents who provided measures of parental interest in education, and this may not be a valid way of measuring such attitudes. It is possible that working class parents have similarly positive attitudes towards education as middle class parents, and want to spend as much time in helping their children, but are restricted by structural constraints. For example, life in general is often more difficult to manage due to lower incomes and poorer housing, which may mean that there is less time for these activities, and they have to be given lower priority. The association reported for being read to and marital status in some ways reflects these constraints in as much as single parents probably do not have as much time to read to their children as two parents have.

It is also possible that working class families feel less confident about helping their children with these sorts of activities than middle class families. If this were true, Tizard *et al.* (1982) demonstrated that this need not be an overwhelming constraint. With the help and encouragement of the research team, who regularly visited parents at home, working class families in the Haringey Project were both willing and able to listen to their children read. It may need this kind of intervention, though, to develop these families' confidence in this area, and without this support they are less inclined to help.

The association between the amount of reading aloud and reading progress was not marked. Our sample was a year younger than that of the Haringey study, which meant that during the main period of our data collection, the youngest children in the sample would have had fewer than three terms at school and many may still have been at the "emergent reader" stage. Maybe parental help is more effective for older children, although the results in Tizard *et al.* do not support this view. It is also possible that the emphasis we put on time in activities was misplaced. We can say nothing about how parents listened to their children read nor how they read to them. Nor do we know what it was the children were reading. These are, of course, important considerations which may affect progress, but to measure these aspects of parental help would have required us either to have tape-recorded sessions (like Hannon *et al.*, 1986), or to have observed within homes using ethnographic methods (e.g., Heath, 1983) or systematic observation. Another consideration is the school's reading policy and its encouragement of parental help. Children's reading activities at home may need to be closely tied to their reading at school for there to be any benefit in terms of reading test scores. Although our data do not reveal any differences between schools for the amount of reading aloud, there were differences in frequency. However, the number of children from each school was small and as we did not collect information about each school's reading policy and practice, we were not able to explore this further.

A more important predictor of reading progress than reading aloud was the amount of time children spent reading on their own, a variable not considered in previous research. We do not know why some children chose to read more than others but it is plausible to suppose that children who start to make good progress in reading for other reasons then choose to give more time to reading on their own. Nevertheless, our results do suggest that future intervention studies might consider not only encouraging parents to hear their children read but also encouraging children to read on their own at home (and schools and parents to provide a range of books for them to read).

The method of data collection used here produced much lower estimates for the frequency of reading aloud than recent research has done (see Plewis *et al.*, 1990, for details). Nevertheless, for the children in our sample, the mean amount of reading at home (aloud plus on own) was 80 minutes per week whereas the mean amount of reading that was observed in year 2 classes at school by Tizard *et al.* was about 40 minutes per week. Clearly children at this age read more at home than they do at school.

ACKNOWLEDGMENTS.—This research was funded by ESRC as part of an extension of its grant to the Thomas Coram Research Unit as a Designated Research Centre. The study was carried out in Inner London Education Authority schools and we would like to thank staff, pupils and parents for their co-operation. We also thank colleagues at the Thomas Coram Research Unit and Jenny Hewison for comments on earlier versions of this paper, one of which was presented to the 1989 British Psychological Society annual conference.

Correspondence and requests for reprints should be addressed to Ian Plewis, Thomas Coram Research Unit, Institute of Education, University of London, 41 Brunswick Square, London WC1N 1AZ.

REFERENCES

CACE (Central Advisory Council for Education) (England) (1967). *Children and Their Primary Schools.* London: HMSO.

Fuller, W. A. (1987). *Measurement Error Models*, New York: Wiley.

Halsey, A. H. (Ed.) (1972). *Educational Priority No. 1: EPA Problems and Policies.* London: HMSO.

Hannon, P. (1987). A study of the effects of parental involvement in the teaching of reading on children's reading test performance. *Br J. educ. Psychol.*, 57, 56-72.

Hannon, P. (1989). How should parental involvement in the teaching of reading be evaluated? *Br. educ. Res. J.*, 15, 33-40.

Hannon, P., Jackson, A., and Weinberger, J. (1986). Parents' and teachers' strategies in hearing young children read. *Research Papers in Education*, 1, 6-25.

Heath, S. B. (1983). *Ways with Words.* New York: Cambridge University Press.

Hewison, J. (1988). The long term effectiveness of parental involvement in reading: a follow-up to the Haringey reading project. *Br. J. educ. Psychol.*, 58, 184-190.

Hewison, J., and Tizard, J. (1980). Parental involvement and reading attainment. *Br. J. educ. Psychol.*, 50, 209-215.

McCullagh, P. (1980). Regression models for ordinal data (with Discussion). *J. Royal Stat. Soc.*, B, 42, 109-142.

Osborn, A. F., and Milbank, J. E. (1987). *The Effects of Early Education.* Oxford: Clarendon Press.

Plewis, I. (1985). *Analysing Change.* Chichester: Wiley.

Plewis, I., Creeser, R., and Mooney, A. (1990). Reliability and validity of time budget data: children's activities outside school. *J. Offic. Stat.* (forthcoming).

Tizard, B., Blatchford P., Burke, J., Farquhar, C., and Plewis, I. (1988). *Young Children at School in the Inner City.* Hove, Sussex: Lawrence Erlbaum.

Tizard, J., Schofield, W. N., and Hewison, J. (1982). Collaboration between teachers and parents in assisting children's reading. *Br. J. educ. Psychol.*, 52, 1-15.

(*Manuscript received 15th October*, 1989)

APPENDIX

The statistical model used to relate 3R activities to progress was a multiple regression model as described in Plewis (1985):

$$y_i = b_0 + b_1 x_i + b_2 z_i + e_i$$

where y_i is the test score at the end of year 1 for individual i, x is the test score at the end of reception and z is the measure of the activity. We are interested in the value of b_2 — the effect of z on y for fixed values of x or, alternatively, the association of z with progress. However, the estimation of the regression coefficients is complicated by the fact that x and, particularly, z are measured with error. Ignoring measurement error in the explanatory variables leads to biased estimates of b_1 and b_2 — see Fuller (1987). Estimates of the reliabilities of z are given in Plewis *et al.* (1990); they ranged from 0.45 for the amount of reading on own to 0.65 for time spent with books and total teaching, with somewhat higher estimates for measures of frequency. Two estimates of the reliability of x were considered, 0.95 and 0.85. The basic equation given above was re-estimated with corrections for measurement error introduced using the methods described by Fuller (1987). The results reported in the text incorporate these corrections.

RESEARCH NOTES

Br. J. educ. Psychol., 60, 338-348, 1990

A NEW CRITERIUM IN THE ASSESSMENT OF THE BEHAVIOUR OF THE DRAWING CHILD

BY PHILIPPE WALLON
(Santé mentale et déviance de l'enfant et de l'adolescent, INSERM U69, Montrouge, France)
AND CLAUDE BAUDOIN
(Unité de Recherches Statistiques, INSERM U21, Villejuif, France)

SUMMARY. A descriptive tool is proposed for the evaluation of the process of the child's spontaneous drawing (i.e., without a model to copy). The process of drawing displays characteristics not visible when the drawing is finished, and so it has a major capacity for discrimination. This method relies on a system for coding the set of lines in the drawing, using an ordinal scale that allows the study of the effects of age throughout the primary school period. This classification may be applied to the drawing of either the human figure or the dog, and allows a comparison between different themes. This method was explored on a sample of 48 primary school children.

INTRODUCTION

Goodnow (1977, p. 59) asked the question: "Does the sequence of drawing in children's pictures help account for the picture drawn?" For her, children's drawings "reveal a great deal of order". She concluded (p. 81) that research on the sequence of drawing needs a variety of shapes, which should be simple and also be uniform. The first solution suggested was to ask people to copy geometric shapes (Gesell and Almes, 1946; Zazzo, 1950; Bender, 1957; Goodnow and Levine, 1973; Lurcat, 1974; Bernbaum, 1974; Blau 1977). This makes the sequence easy to observe but cuts down the variety of the drawing.

It is well known, at least since Goodenough (1926), that the evolution of spontaneous drawing is related to the child's psycho-affective maturation (Freeman and Cox, 1985). Very few studies assess the process of spontaneous drawing (Goodnow, 1977; Freeman, 1980; Van Sommers, 1984), but none deals with the developmental aspect. In scribbling and the very first stage of spontaneous drawing of the human figure, Eng (1954), Prudhommeau (1947) and Goodnow (1977) described the beginning of the drawing, the sequence of lines, the rotations and the alterations from right to left. The literature shows a positive correlation between the counterclockwise direction and the level of schooling. The "clockwise" circle ("Torque", Blau, 1977) could be the sign of a dual cerebral dominance; this was criticised by Jarman and Nelson (1980) and Furlong (1981). Goodnow and Friedman (1972) gave the child special orders for drawing so that variations in sequences could be highlighted. However, these different studies are not able to provide the elements which would permit following the child's graphic maturation, in the field of educational psychology.

We propose a descriptive tool for assessing the sequence of lines in a primary school child's spontaneous drawing. We use two themes: the human figure and the dog. Descriptive criteria may be created in order to provide an interpretation based on logical operations that may be computerised, without reference to psychometric or projective psychology. Forty-eight primary school children were studied. As Goodnow noted (1977, p. 22, citing Lashley): ". . . the general point that understanding the nature of serial action was a key to understanding the nature of skill".

METHOD

Procedures: The analysis of the "executing scheme"

The drawing itself is not analysed directly. Instead, an "executing scheme" is produced by an observer during the child's performance (Wallon, 1987). The chronology of the execution is carefully recorded: the starting point (SP) of the drawing and the finishing point (FP), the lines sequenced with a chronological number, the beginning point (BP) of each line, its end point (EP), its direction and the relative positions of the set of lines. Neither the proportions nor the size of the different elements of the drawing are taken into consideration. Four drawings are shown (Figures 1 to 4), which are not necessarily representative of the sample studied, but illustrate various types among the 96 collected drawings.

Basic items

The scale for the coding of each drawing rests on two main elements: (i) the starting point and (ii) the set of lines, which are differentiated for the head and the trunk. Our classification does not distinguish the limbs because they have a dynamic in common with the trunk. Figure 5 illustrates the basic structures for coding the line, and Figure 6 shows the scheme of the different items coded. The first rubric (Figure 6) codes the starting point (SP). Coding differs for the human figure and the dog. For the human figure, the coding scheme for the SP used six items, along a vertical axis. The different items may be grouped according to this axis: Superior: head + neck; Middle: upper limbs + lower trunk; Inferior: lower limbs + feet. The coding of the starting point for the dog uses eight items. They may be grouped according to a superior-inferior axis: Superior: head + back + tail; Middle: throat + buttocks; Inferior: front legs + between legs + rear legs; or according to a cephalo-caudal axis: Cephalic: head + throat; Middle: back + front legs + between legs; Caudal: rear legs + buttocks + tail.

The second schema (Figure 6) deal with the set of lines that concerns the trunk and head. Coding is done on the basis of a scale that has been sub-divided into eight items and numbered according to an increasing degree of complexity. These items are grouped by branch. The absence of the trunk corresponds to two completely different types of drawings; either primitive (for example the classic "head-trunk") or very advanced (the "face"). In between, we have set up an "ordinal" scale, that ranges from the simplest drawing style to the more complex. The two simplest lines are represented (Figure 5) by the "accumulated" line and the stick-line, and these exist only for the trunk. The "accumulated" line is composed of a stack of curvilinear lines whose beginning and end are tied to the previous lines. Among the "organised" forms, the "stick" represents the unidimensional aspect of the line, in contrast to the bidimensional aspect delimiting a "surface". The stick-line is classified more advanced than the accumulated line because the end remains free and attests to a beginning of organisation in the drawing. The simplest surface is a "loop", defined by an uninterrupted line, no matter what its shape, where the beginning and the end meet. Otherwise the line is viewed as "open". Open lines may be "simple" or "complex". They are "simple" when the lines immediately follow one another (the beginning of a new line is tied to the end of the previous one) and all the lines are going in the same rotation; they are "complex" when the lines are going in different rotation. In the latter case, according to the axis of the drawing, associating both the starting and finishing point, the succession of the lines may be "continuous" on the same side, or "alternate" on both sides. The underlying hypothesis is that for any given child, advancement along the scale from "accumulated" to "stick" or "loop", to "continuous complex" and finally to "alternate complex" will occur over time. This scale is completed with nominal items (Figure 6). The line is noted as "upward", "downward" or "horizontal" according to the orientation of the axis of the drawing. The orientation for "loop" or "open simple" lines, is defined as "clockwise" or "counterclockwise".

Notation

Each child's drawing is coded by three numbers as shown in Figure 6. The first one indicates the mode of the starting point, the second describes the head line and the third the line of the trunk. The code of the starting point is a single number. The head and the trunk are each coded by a two-digit number: the notation of the eight classes of the "ordinal" variable goes into the units column; the "nominal" variable takes the tens column as indicated in Figure 6. The digit "0" indicates that data are missing and "9" denotes "not applicable". In the case of the human figure, the starting point SP has six classes, the dog, eight. The lines of the head are described according to 11 classes, and the trunk, 18. Consequently, a head made of an open simple, clockwise line would be coded 15, a trunk with an alternate complex line, downward, would be coded 17 (see other examples in Figures 1 to 4).

Sample and Conditions of Experiment

The study was conducted in Montreal, Canada, among normal, primary school children. Forty-eight children (24 boys), mean age 10 years (SD = 2, range 6.5 to 13 years), were randomly selected from four schools, stratified on the basis of administrative area of the Montreal Catholic School Commission (MCSC). Within each school, a boy and a girl were randomly selected from each of the six grade levels, 1st to 6th. Each choice was submitted for approval to the pedagogic board of the school. Children were required to draw two subjects: a human figure and a dog. The performing order was randomised. The only instructions given were to "do the best that you can".

FIGURE 1

(Girl, 6 years 9 months old). The drawing starts at the top of the head with an uninterrupted (number 1), closed, rounded, counterclockwise line. The neck is made up of two lines, almost parallel (2 and 3). The dress is delimited by two lines (4 and 5), unequal in length, complex as they define the beginning of the sleeves; they both start from the low neck and join on the sides, at the bottom of the dress. From the bottom of the drawing, the child comes back to the top with the hair, composed of three lines (6, 7, 8), forming a rounded closed line together. We note here the alternate aspect of the lines 6 and 7. The details of the face (9) are not taken into consideration. The four limbs look like a "stick". The lower limbs (10, 11) were drawn before the upper limbs (14, 16). Feet (12, 13) and hands (15, 17) are drawn with a counterclockwise loop line.
CODE: beginning: head (1), head: counterclockwise loop-line (24), trunk: downward alternate complex line (17).

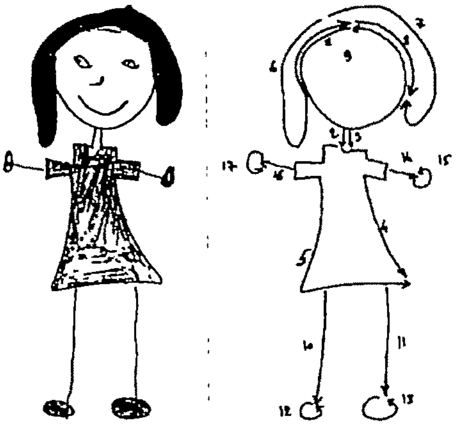
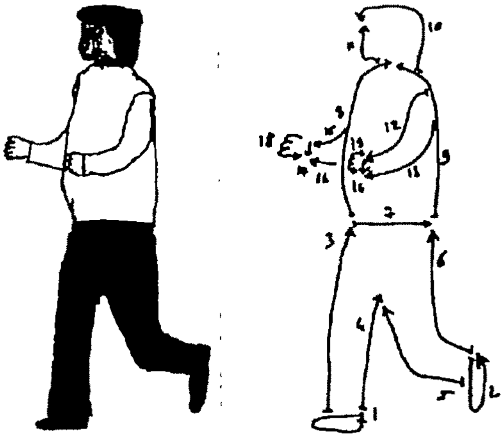


FIGURE 2

(Boy, 12 years 2 months old). Coding reveals the upward and alternate drawing of all the lines; the drawing starts from the feet and ends with the hands. The two feet are made of counterclockwise loop-lines (1 and 2). From here, four upward lines (3, 4, 5, 6), from left to right, delimit the lower limbs, ending at the waist, which closes the entire figure. Again, from left to right, two upward and symmetrical lines close the trunk. Similarly, the head is made up of two upward lines, that are asymmetric from left to right. The building of the arms follows the alternate rule: for each of them, two parallel lines start from the basis (12, 13 and 15, 16) and are closed at the wrist by a small line (14, 17). Drawing of the hands does not immediately follow that of the arms. The hands end the drawing (18, 19).
CODE: beginning: feet (6), head: upward alternate complex line (27), trunk: upward alternate complex line (27).



The drawing took place during schooltime, in a classroom. Observation of the execution of the drawing was done by the same observer for the entire duration of the study. Each child was observed individually. The observer stood approximately one metre behind the child, on the left or the right according to the hand the child used. During the session, the observer reproduced the drawing, using the basic items described in Figure 5. Later, the entire scheme was coded by the researcher according to three rubrics (Starting Point, Head Line, Trunk Line), using the scale presented in Figure 6.

RESULTS

Table 1 shows the frequency of the different modes for beginning the drawing of a human figure or a dog, and their cross-tabulation. A majority of children started by drawing the head (N=38, 79 per cent for the human figure and N=31, 65 per cent for the dog). One out of every two children (N=26, 54 per cent) began both themes at the head. Among the remaining rubrics, no clear "second pole" was observed. When grouping the three items, 38 (79 per cent) children started drawing the human figure at the upper part, seven (15 per cent) at the lower part and three (6 per cent) in the middle; for the dog, the distribution was 40 (83 per cent), six (13 per cent) and two (4 per cent), respectively. When the second classification for the dog was used, 32 (67 per cent) children started by its cephalic part, seven (15 per cent) by its caudal part and nine (19 per cent) in the middle. When the data for the human figure and the dog were cross-tabulated, 33 (69 per cent) children began at the upper part and three (6 per cent) at the lower part.

TABLE 1

Starting Point of the Drawing. Comparison "Between-Theme", Man and Dog

		MAN						
		Head	Neck	Upper limbs	Lower trunk	Lower limbs	Feet	Total
	Head	26	0	1	1	3	0	31
	Throat	1	0	0	0	0	0	1
	Front legs	0	0	0	0	1	1	2
DOG	Between legs	2	0	0	0	0	0	2
	Rear legs	1	0	0	0	1	0	2
	Buttocks	1	0	0	0	0	0	1
	Tail	4	0	0	0	0	0	4
	Back	3	0	1	0	0	1	5
	Total	38	0	2	1	5	2	48

The human figure and the dog were cross-tabulated according to different modalities for the head (Table 2a) and the trunk (Table 2b). The line of the human's head (Table 2a) was subdivided into two groups: 36 (75 per cent) were composed of a "loop" and 11 (23 per cent) others were made of "complex alternate" lines. Only one drawing was "open simple". Similar bipolarity was not found in the drawing of the dog's head. This was more regularly distributed, with a majority of heads designated as "loop" (N=20, 42 per cent), 14 (29 per cent) as "open simple" drawings and 14 (29 per cent) as "open complex". The "open complex" were either "continuous" (N=4, 8 per cent) or "alternate" (N=10, 21 per cent). The line of the trunk of the human figure (Table 2b) showed a strong polarity for the type "complex alternate" (N=29, 60 per cent). The remaining drawings were distributed in small numbers into the other classes. The trunk of the dog was distributed more regularly among the different categories with 18 (38 per cent) designated as "complex alternate"drawings, 10 (21 per cent) as "loop", nine (19 per cent) as "simple" drawings and nine (19 per cent) as "continuous". Three children drew the human figure as a "face" (with no trunk), and one of them drew a trunkless dog as a "face".

It is possible to summarise the results in Tables 2a and 2b by considering the transition in complexity from the drawing of the human figure to that of the dog (Table 2c). This allows the definition of a status of "progression", "stability" or "regression" with reference to the ordinal part of the scale. In Table 2a, 19 drawings (40 per cent) fell under the main diagonal

FIGURE 3

(Girl, 9 years 10 months old). The drawing of this animal is "upward". It starts at the lower limbs and is drawn successively from front to rear (1, 2, 3, 4), and ending at the head. The limbs are drawn with a stop-start again; the end of line 1 is free because the beginning of line 2 is not tied to it. This is also true for the ends of lines 3 and 4. The back drawing of the last hind leg (4) diverges to make the tail. The rest of the line, for the entire trunk-head, is closed with four alternate lines between rear and front (5 and 6, 7 and 8).
CODE: beginning: between legs (3), head: upward alternate complex line (27), trunk: upward alternate complex line (27).

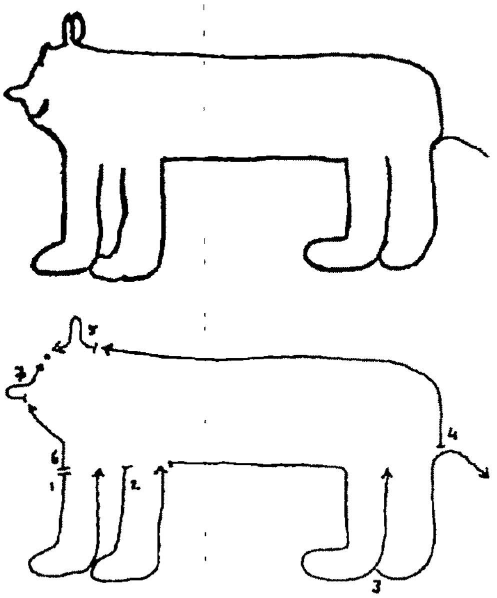
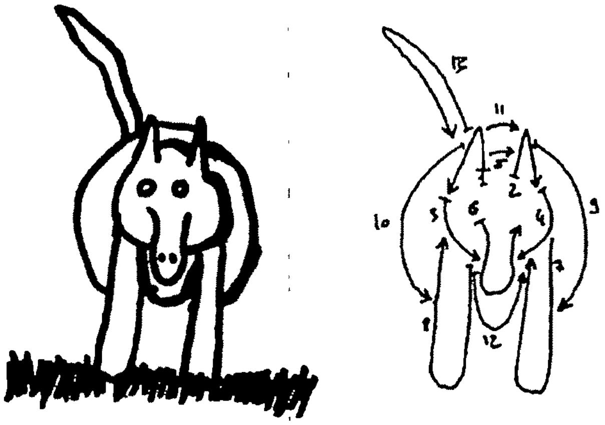


FIGURE 4

(Boy, 11 years 5 months old.) This is the drawing of an animal, characterised by its frontal perspective. The different elements of the drawing are added to each other in a centrifugal way. The lines are drawn alternately on both sides of the vertical median axis. The drawing starts at the head, precisely at the ears (1, 2), then goes alternately from right to left with the cheeks (3, 4), the legs (tied to the head, 7 and 8), and the sides (9 and 10).
CODE: beginning: head (1), head: downward alternate complex line (17), trunk: downward alternate complex line (27).



indicating a progression in the drawing of the dog's head in comparison with that of the human. Four (8 per cent) drawings fell above this diagonal, showing a regression. In the case of the trunk (Table 2b), 20 drawings (42 per cent) indicated a regression in complexity in the drawing of the human figure compared to that of the dog, and six showed a progression. Table 2c compares the different states of transition between the human figure and the dog. This showed a stability between both head and trunk items in 15 children (31 per cent). In addition eight cases showed a stability in the drawing of the head and a regression of the dog's trunk as compared to the human, and five cases (10 per cent) showed a stability in the trunk and a "progression" in the dog's head as compared to the human. In 10 children (20 per cent), the drawing of the dog's head progressed in comparison to the human figure, whereas the trunk regressed. Two children (4 per cent) drew a dog which regressed globally (head and trunk); on the other hand, four children (8 per cent) displayed a dog that was globally more elaborate than the human.

TABLE 2

Comparison Between Drawing of the Man and of the Dog

TABLE 2a: The Head
(Lo=Loop, Si=Simple, Co=Continuous, Al=Alternate)

		MAN				
		Lo	Si	Co	Al	Total
DOG	Lo	18	0	0	2	20
	Si	12	0	0	2	14
	Co	3	1	0	0	4
	Al	3	0	0	7	10
	Total	36	1	0	11	48

TABLE 2b: The Trunk
(Ha=Head alone, Ac=Accumulated, St=Stick, Lo=Loop,
Si=Simple, Co=Continuous, Al=Alternate, Fa=Face)

		MAN								
		Ha	Ac	St	Lo	Si	Co	Al	Fa	Total
DOG	Ha	0	0	0	0	0	0	0	0	0
	Ac	0	0	0	0	0	0	0	0	0
	St	0	0	1	0	0	0	0	0	1
	Lo	0	0	1	4	1	1	3	0	10
	Si	0	0	0	0	2	0	7	0	9
	Co	0	0	0	0	1	1	6	1	9
	Al	0	2	0	0	2	0	13	1	18
	Fa	0	0	0	0	0	0	0	1	1
	Total	0	2	2	4	6	2	29	3	48

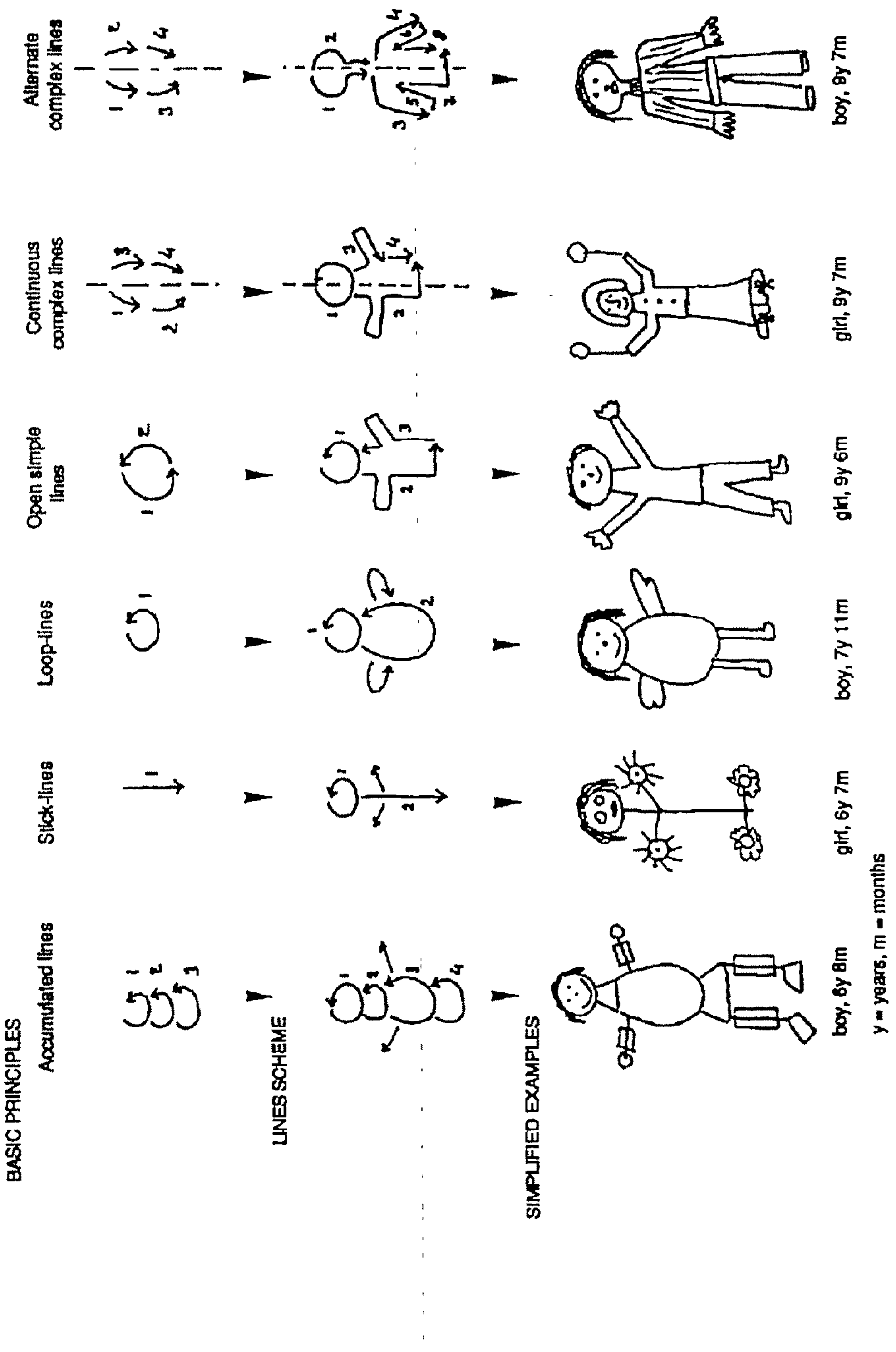
TABLE 2c: Transition of the Drawing of the Dog Compared to the Man

		Transition of the TRUNK from the MAN to the DOG			
		regression	stable	progression	total
Transition of the HEAD from the MAN to the DOG	regression	2	2	0	4
	stable	8	15	2	25
	progression	10	5	4	19
	total	20	22	6	48

The "ordinal" character of the coding allows the establishment of a scale of progression-regression, from the data in Tables 2a and 2b. A progression in the drawing was defined whenever the child goes, from the man to the dog, to a code of a superior rank. Symmetrically, a regression in the drawing was defined when the child moved to a code of inferior rank. Stability of the drawings is defined by unchanged codes.

FIGURE 5

THE BASIC SHAPES FOR CODING THE LINE



y = years, m = months

TABLE 3

Transition of the Head Compared to the Trunk

		DOG			
		regression	stable	progression	total
MAN	regression	14	15	3	32
	stable	5	10	1	16
	progression	0	0	0	0
	total	19	25	4	48

Classes of 'progression', 'stability' and 'regression' were defined by the comments in Table 2c.

Table 3 assesses the head and the trunk according to the same theme. Regarding the trunk, the three categories "cumulated", "stick" and "loop" were collapsed so they could, for the head, refer to the basic category "loop". For the human figure, two-thirds of the drawings (N=32, 67 per cent) showed a regression in the head when compared to the trunk, and for the last third of the drawings (N=16, 33 per cent), the complexity of the head drawing was similar to that of the trunk. Among these 16 cases, however, four heads designated as "loop" were associated with a trunk defined as "cumulated" (N=2, 5 per cent) or as "stick" (N=2, 5 per cent). No progression in the head drawing was observed when it was compared to the trunk. For the dog, regression in the head as compared to the trunk was less frequent (N=19, 40 per cent), leading to a greater stability (N=25, 52 per cent). Four drawings (8 per cent) of the head were designated as "progression" when the head was compared to the trunk, showing, somehow, very similar lines (two drawings had a "complex-continuous" head with an "open-simple" or "complex-alternate" trunk, one had an "open-simple" head and a "loop" trunk, and the last, a "complex-alternate" head with a "continuous-complex" trunk). When the two themes were crossed, the same type of transition from the trunk to the head was observed for dog and man in 24 children (50 per cent, on main diagonal). For the other cases, dissymmetry between the human figure and the dog was clearly apparent. Among the 25 drawings of dogs that showed a similar line for the head and the trunk, 15 (60 per cent) presented some regression in the human's head when it was compared to the corresponding trunk. In addition, three out of the four drawings designated as "progression" in terms of the dog's head as compared to trunk, also displayed "regression" in the character's head when it was compared to the trunk.

Table 4 shows the relation between the age of the child and the ordinal part of the scale. Line complexity increased with age. No relationship was evident between the starting point of the drawing and age.

TABLE 4

Correlation Between Age of the Child
(in Number of Days) and Coding of the Line

THEME		Correlation coefficient*	P <
MAN	Head	0.491	0.001
	Trunk	0.470	0.001
DOG	Head	0.408	0.003
	Trunk	0.341	0.010

* Spearman correlation coefficient

DISCUSSION

Our work displays a method for coding and analysing the execution of a definite theme drawing. Our method was explored on a small sample of children which was stratified according to school and class, in order to have a representative sample of different ages and socio-cultural backgrounds. We have shown, using two themes (the human figure and the dog), discrimination at different levels of complexity: (i) for the same theme, the trunk was a more complex set of lines

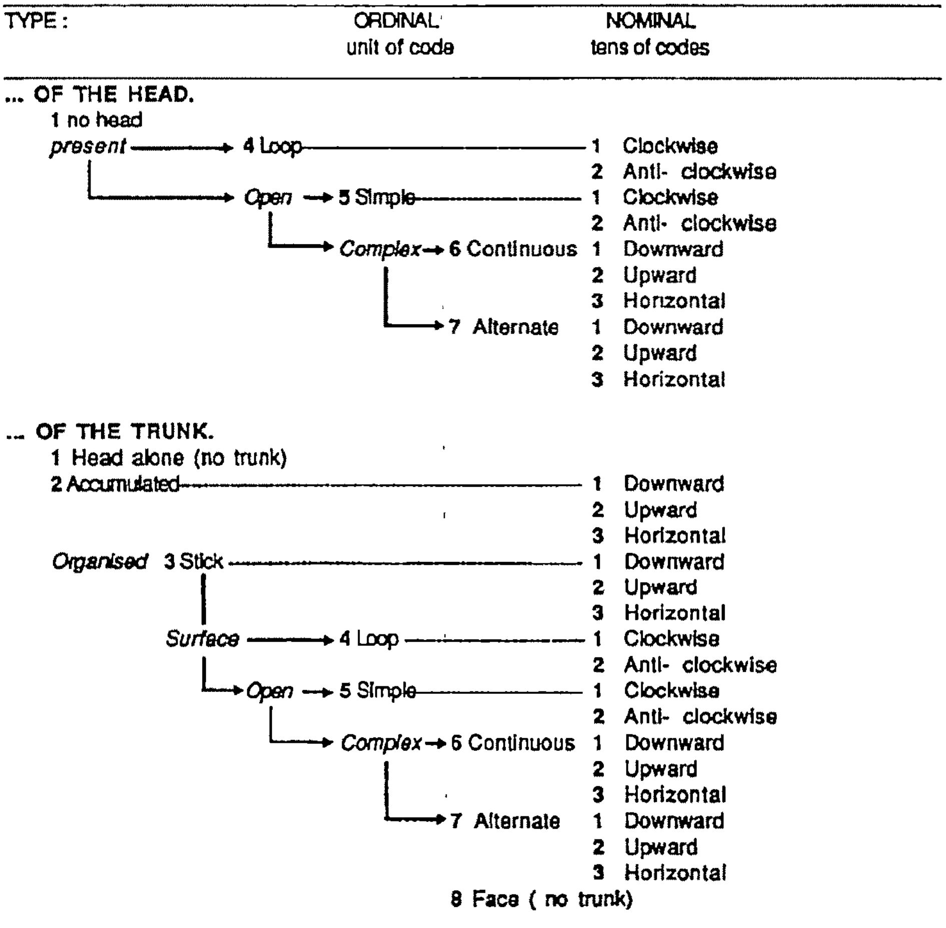
FIGURE 6

CODING GRID OF THE STARTING POINT AND THE LINES OF THE DRAWING

CODING OF THE STARTING POINT.

	MAN		DOG
1	Head	1	Head
2	Neck	2	Throat
3	Upper limbs	3	Front legs
4	Lower trunk	4	Between legs
5	Lower limbs	5	Rear legs
6	Feet	6	Buttocks
		7	Tail
		8	Back

CODING OF THE LINES.



0=missing data 9=not applicable

than the head, (ii) the dog's head was a more complex set of lines than that of the human figure, (iii) the human's trunk was more complex than that of the dog, (iv) the human presented a strong polarity for both head and trunk (simple head, complex trunk), whereas the line of the dog was more modulated into the different groups of our classification, (v) the level of complexity was always related to the age of child and the correlation is always positive.

The difference in the evolution of the set of lines, from the human to the dog, may be due to three reasons: (i) in the human figure, both head and trunk may constitute the two poles, of which the evolution was almost independent. On the other hand, the drawing of the dog appeared with more homogeneity, since the head often displays less individuality; (ii) the line "in loop", considered as primitive, is best adapted to the oval of the face: that could explain why the head of the human figure remained at an early stage of evolution whereas the head of the dog progressed; (iii) the trunk of the animal, relatively simple to draw, might not tax all of the child's abilities. With regard to the human figure, the trunk, adorned with clothes, with its vertical symmetry, may allow the child to express his ability for complex drawings (continuous or alternate).

We chose to limit our study to the line, for two complementary reasons: (i) its execution, both in space and time, has almost infinite possibilities of variation which are, undoubtedly, not merely due to a random effect, (ii) such great variations do not imply a large heterogeneity in the analysis.

The "Starting Point" refers to the main element of the theme (head for the man and the dog). Some children started their drawing in an unusual spot (for instance: the feet), and performed the whole drawing from this spot. This raised the question of their being able to anticipate the drawing.

The coding of the lines was based on empirical hypotheses elaborated from observations of children in the course of drawing from a previous work (Wallon, 1987). This work resumed and completed the Lowenfeld's (1957) classification (additive and organic drawings), and classified the lines of the finished drawing according to three types of lines: additive, embracing (from Goodnow, 1977) and stop-start lines (the line stops in the course of the drawing and starts somewhere else, this new line joins the first one, without gap or crossing — Wallon, 1987). In the present paper, we have completed our classification, so that it may be applied to the drawing process postulating that it is performed according to a chronologic sequence: (i) first two primitive lines (the "stick" and the "accumulated"), (ii) "loop" lines. (i) and (ii) corresponding to our additive lines and Lowenfeld's (1957) geometrical lines. Then more complex lines correspond to our embracing and stop-start lines: (iii) lines all in one piece, the child draws a line without lifting his pencil, (iv) closed simple lines, the child lifts his pencil but comes back to the line where he left it moving in the same direction, (v) then complex lines, continuous and alternate. It is impossible, at present, to assert whether a succession exists between the two complex lines. The child sometimes draws in an alternate way according to the symmetry axis of the drawing, especially in the human figure. In the case of the dog, the axis of symmetry is not so obvious, and the alternation is less marked. It is important to remark that (iii) to (v) are indiscernible on the finished drawing, where they are only observed as "embracing lines"; our classification displays a hidden evolution on the finished drawing. In addition, the type "complex alternate" could lead to two types of finished drawing: embracing lines and stop-start lines.

Our work, in contrast to that of Van Sommers (1984), who made a general analysis of the graphic execution, deals only with drawings on an imposed theme. Our themes are limited to two specific ones (susceptible to a "semantic contamination", according to Van Sommers). At this stage, our work is an exploratory approach and can not encompass the variety of the sequence of drawing. However, it shows it is possible to study the sequence of lines in the child's spontaneous drawing.

The small sample size, and also the method of the experiment, includes constraints liable to introduce biases: using only a felt-tip pen, on a horizontal table, in a school room. The method for collecting the drawing does not exactly reflect its process: when building the scheme of execution, the speed or the rhythm of each line is not noted. Finally, the method of analysis is also limited: the coding grid groups rather different lines and does not differentiate the direction of the lines or the alternation in relation to the left or right axis (this is important for the right or left-handed child).

Our method provides a description of the drawing, without any immediate psychological interpretation, whereas the research in the literature was finalist because the interpretation of the drawing, in terms of level of development (Goodenough, 1926; Harris, 1963) or personality (Koch, 1949), is without any independent descriptive stage.

Our work has tried to define a path of investigation which relies on the process of drawing and not on the representation of a model. Most authors look principally at the resemblance of the drawing to a model. They insist on the "realism" or "inadequacy" of such and such an item of the drawing (Goodenough, 1926; Machover, 1953; Harris, 1963).

Our analysis concerns the dynamics of the execution which, just as the static items of the finalised drawing, compose the specific "signature" of the child. It is quantitative and concerns only drawings of an imposed theme, and belongs to the field of educational psychology. Most of the studies are based on the drawing when it is finished. The few studies concerning the child in the course of drawing may be classified into two groups: (i) quantitative studies performed on populations of children, concerning mostly geometric figures (circles, cycloids, etc) or (ii) drawings performed after a theme, mainly the human figure. Such studies belong to the field of differential rather than educational psychology, as their method focused on primitive drawings, qualitative or individual studies, concerning the way the child draws, i.e., not only the line but also the child's behaviour (talking, comments on the drawing). They often concern young scribbling children, observed through film or video recording. Some clinical cases also exist, which do not suggest any methodological tools (drawings performed in the course of psychotherapy).

Our methodology allows a joint analysis of both the human figure and the dog. From a theoretical viewpoint, this evaluation of the dynamics should provide a connection between the child's drawing and his global activities. The literature does not refer to any studies that provide a joint analysis of several themes. The House-Tree-Person (H.T.P., Buck, 1948) gathers three independent analyses, i.e., the characters specific to a theme are not valid for another, at least in the details. The available studies account only for the themes man, tree and house, and not for the drawing of an animal, which remains a relatively unexplored theme.

Acknowledgements.—This work was supported by a grant from CNAMTS-INSERM (International Research Project no. 993347). We thank Montreal Catholic School Commission and the directors of the schools for their assistance, and B. Balkau and J. Chwalow for reviewing the manuscript.

Correspondence and requests for reprints should be addressed to M. Philippe Wallon, Unité de Recherches Statistiques, 16 Av. Paul-Vaillant-Couturier, 94807 Villejuif Cedex, France.

REFERENCES

Bender, L. (1957). *Un Test Visuo-moteur*. Paris: Presses Universitaires de France.
Bernbaum, M. (1974). Accuracy in children's copying: the role of different stroke sequences and school experience. PhD. Dissertation, George Washington University.
Blau, T. H. (1977). Torque and schizophrenic vulnerability. *Am Psychol.*, 12, 997-1005.
Buck, J. (1948). The H.T.P. Test *J. clin. Psychol.*, 4.2.
Eng, H. (1954). *The Psychology of Children's Drawings* (2nd ed.) London: Kegan Paul.
Freeman, N. H. (1980). *Strategies of Representation in Young Children: Analysis of Spatial Skills and Drawing Processes*. London: Academic Press.
Freeman, N. H., and Cox, M. V. (Eds.) (1985). *Visual Order*. Cambridge: Cambridge University Press.
Furlong, M. J. (1981). Torque: an at-risk indicator of reading behaviour problems. *J. clin. Psychol.*, 3, 165-167.
Gesell, A., and Almes, J. (1946). The development of directionality in drawings. *J. genet. Psychol.*, 68, 45-61.
Goodenough, F. L. (1926). *Measurement of Intelligence by Drawings*. Chicago: World Books.
Goodnow, J. (1977). *Children's Drawing*. Glasgow: Fontana/Open Books.
Goodnow, J. J., and Friedman, S. (1972). Orientation in children's human figure drawings: an aspect of graphic language. *Dev. Psychol.*, 254, 416-417.
Goodnow, J. J. and Levine, R. (1973). The grammar of action: sequence and syntax in children's copying of simple shapes. *Cognit. Psychol.* 4, 82-98.
Harris, D. B. (1963). *Children's Drawings as Measures of Intellectual Maturity*. New York: Harcourt, Brace and World.
Jarman, R. F., and Nelson, J. G. (1980). Torque and cognitive ability: some contradictions to Blau's proposal. *J. clin. Psychol.*, 2, 458-464.
Koch, Ch. (1949). *Der Baumtest.* Berne: Huber.
Lowenfeld, V. (1957). *Creative and Mental Growth.* (3rd ed.). New York: Macmillan.
Lurcat, L. (1974). *Etudes de l'acte graphique*. Paris: Mouton.
Machover, K. (1953). *Personality Projection in the Drawing of the Human Figure* (2nd ed.). Springfield, Ill.: Thomas.
Prudhommeau, M. (1947). *Le Dessin chez l'Enfant*. Paris: Presses Universitaires de France.
Van Sommers, P. (1984). *Drawing and Cognition.* Cambridge: Cambridge University Press.
Wallon, P. (1987). *Le Dessin Spontané d'Animaux chez l'Enfant*. Brussells: Editest.
Zazzo R. (1950). Le geste graphique et la structuration de l'espace. *Enfance*. 3-4, 204-220.

(*Manuscript received 13th February*, 1989)

Br. J. educ. Psychol., 60, 349-355, 1990

PARENTAL INVOLVEMENT IN THE TEACHING OF READING: A COMPARISON OF HEARING READING, PAIRED READING, PAUSE, PROMPT, PRAISE, AND DIRECT INSTRUCTION METHODS

By DAVID J. LEACH
(Murdoch University, Western Australia)
AND SUSAN W. SIDDALL
(Ministry of Education, Western Australia)

SUMMARY. The parents of 40 children in two grade 1 classrooms were randomly assigned to receive brief training in one of four instructional methods for helping their child to read or to hear their children read at home. The tutoring methods were Hearing Reading, Paired Reading, Pause, Prompt, Praise, and Direct Instruction. The results showed that the use of the additional instructional strategies included in the Direct Instruction and Paired Reading tutoring methods led to faster progress by the children receiving them than by children whose parents simply heard them read. Some theoretical and practical implications of the study are discussed.

INTRODUCTION

Since 1970, a growing number of publications have shown how parents can be guided to help their children acquire reading skills at home. Such additional help has often led to faster rates of progress than those achieved by schooling alone. Whereas early papers concentrated on parents' use of general strategies for enriching the home "reading environment" and on the effects of enhanced home-school links (Quisenberry *et al.*, 1977), later ones have focused to a greater extent on the effects of formally taught, more complex parent tutoring skills (Topping and Wolfendale, 1985).

A number of strategies and techniques have emerged from this effort which have gained substantial experimental support. At the simplest, least intrusive level of involvement, a consistent finding, with one or two exceptions (e.g., Hannon, 1987), has been that parents who are guided by professionals to listen to their children read at home on a regular basis are more likely to have children with higher reading attainments than parents who do not listen to their children read. This relationship seems to be causative and holds true irrespective of differences between children, such as IQ, and differences in maternal language and child rearing styles (Hewison and Tizard, 1980; Tizard *et al.*, 1982). At a more complex, more intrusive level of involvement, the strongest support seems to have accrued for two parent-delivered tutoring packages — Paired Reading, a method employing simultaneous parent-child reading of texts (e.g. Morgan and Lyon, 1979; Bush, 1983; Robson *et al.*, 1984) and Pause, Prompt, Praise, a method to teach self-correction responses to errors in reading using syntactic and semantic cues (e.g., McNaughton *et al.*, 1981; Love and Van Biervliet, 1984).

Although these methods have been shown to enhance the reading skills of children across ages, abilities and degrees of reading difficulty when compared to no treatment controls, the variety of populations found in evaluation projects has made it impossible to compare their relative effectiveness. Nor has it been possible to weight the contribution of their components, such as practice and motivation on the one hand, and the additional, more time-consuming and demanding strategies of the parent tutoring packages on the other. Despite the promise of the literature, there is need for further investigation of whether guided Hearing Reading at home is all that is needed of parents in order to enhance their children's reading attainment, or whether there are greater gains to be made by investing time and resources in more elaborate parent training programmes.

This study was designed as an initial attempt to compare the relative effectiveness of the Paired Reading and Pause, Prompt, Praise tutoring methods with a basic Hearing Reading strategy when used by parents. We also added a fourth method which lies at the more complex, more intrusive end of the continuum. This was Direct Instruction reading — a comprehensive, task-analysed, phonic-based instructional programme which incorporates fully-scripted lessons and finely graded reading tasks based on a theory of instructional design developed by Englemann and his colleagues (Englemann, 1980). Direct Instruction reading programmes, at least when used by teachers, have received broad research support in the USA (Becker *et al.*, 1980), Australia (Lockery and Maggs, 1982), in the UK (Gregory *et al.*, 1982; Branwhite, 1983). In the main, however, researchers have concentrated on their use with disadvantaged children, low achievers, or children with intellectual disabilities. Surprisingly few (e.g. Noon and Maggs, 1980) have evaluated their use with regular students. Also, as far as can be ascertained, there have been no published evaluations of their effectiveness when delivered solely by parents. Despite these omissions, our own (as yet unpublished) work has led us to believe that average to above average students can make exceptional progress on Direct Instruction programmes, and that parents can be taught how to use them with brief training and are able to achieve equal (if not better) results than those achieved by professional teachers.

It was hypothesised that Direct Instruction, the most comprehensive instructional package, would increase beginning reading skills to a greater extent than either the PR (Paired Reading) or PPP (Pause, Prompt, Praise) methods and that, because of their additional instructional components, the Direct Instruction, Paired Reading, and Pause, Prompt, Praise methods would each be more effective than the minimal, Hearing Reading strategy.

The study was also of some theoretical interest. Firstly, it might be argued that any successful parent involvement in reading could be due to the extra time that was being spent on reading practice plus an increase in their children's motivation. To control for this, all children were to receive the same amount of parental involvement and interest over the experimental period and were to be drawn from the same school classes, thereby increasing the likelihood that any differences found between methods were due to the methods themselves.

Secondly, there has been some criticism in the literature about the lack of comparison groups receiving equally valid, equally novel programmes in the evaluation of Direct Instruction, particularly in the UK (Wheldall and Glynn, 1988). Although this is certainly not true of at least one major evaluation of Direct Instruction conducted in the USA (e.g., Becker, 1977; Becker and Gersten, 1982), the same point could be made about the majority of evaluations of novel approaches to the teaching of reading. In this study, therefore, all tutoring methods were novel for the children concerned.

Thirdly, there has been a continuing debate about the role of errors in learning to read and the optimum way to deal with them in the instructional context. One view is that errors play a central, productive role in that children learn to be more self-regulated readers if they are helped and prompted to obtain appropriate information from them (Clay, 1979; McNaughton, 1988). The opposing view is that, in order to facilitate retention and increase rates of learning, errors should be minimised by programming for high success using graded discrimination, massed examples, preteaching, prompts and immediate corrective feedback (Englemann, 1980). The Pause, Prompt, Praise method utilised here is based on the productive view of errors, whereas the Direct Instruction approach, through its programme design features, plans to minimise errors. The Paired Reading method falls between the two in that it is not programmed to prevent errors but that, when errors occur, immediate corrective feedback is given without prompts to gather information from them.

METHOD

Sample

Parents and their children from two grade 1 classes in two primary schools participated in the study. With the full support of the school staff, all parents in the two classes were asked by letter whether they would like to become involved in a project to teach them how to help with their children's reading at home. Parents of children who could already read at a grade 1 level or above and of children considered by their teachers to have marked learning difficulties were included in the training but their children were excluded from the experimental sample. From the remaining pool of parents, 40 were randomly drawn and then randomly assigned to receive one of the four tutoring methods — Direct Instruction (DI), Paired Reading (PR), Pause, Prompt, Praise (PPP) or Hearing Reading (HR). None of the parents was aware of the overall experimental design.

The final sample of children was composed of 26 boys and 14 girls with chronological ages ranging from 5 years 3 months to 6 years 4 months (M = 5 years 7 months). No child was considered by the teachers to have significant learning difficulties. One girl had a mild speech impediment and one family spoke English as a second language. The children were all beginning readers in the middle of their first term in school. Each made more than 16 errors on the first story of Neale Analysis of Reading Ability. They continued to receive instruction in reading according to the school's normal practices throughout the intervention period.

Parent training

Parent training was carried out at school by two psychologists, one of whom was the second author. Separate training sessions were held for each tutoring group but no formal training was given to the Hearing Reading group. The latter received written guidelines and so could be described as a minimally guided hearing reading group.

All reading sessions were conducted by parents in their own homes for a period of 10 to 15 minutes per weekday for 10 weeks after the completion of the last training group. Three of the designated parent tutors were fathers, one was an 18-year-old sister who spoke the most fluent English in the family, and the rest were mothers. They received one home visit to observe a reading session and a telephone call to check on their appropriate use of the methods taught.

(a) *Hearing Reading (HR)*

The HR group received a brief summary of suggestions for helping with reading at home. This included advice on choosing a suitable time and place for hearing their children read from books brought home from school, talking about the story to be read, giving encouragement and avoiding criticism, and allowing the children time to think about a word when an error occurred before trying to help or correct. Teachers at the school ensured that the children took home appropriate grade 1 level books of their choice after each school day.

(b) *Paired Reading (PR)*

The PR group attended a formal training evening of one and a half hours duration. The training session included:

An introduction to the concept of Paired Reading and an account and demonstration of its two components — simultaneous and individual reading.
A demonstration of how to do Paired Reading.
Supervised role-playing by parents using Paired Reading.

Parents were given a summary of the techniques involved in Paired Reading to take home for reference. The teachers ensured that the children took home a grade 1 book of their choice after each school day.

(c) *Pause, Prompt, Praise (PPP)*

The PPP group also attended a training evening of one and a half hours duration. This session included:

An introduction to the concept of Pause, Prompt, Praise and an account and demonstration of its treatment of errors.
A demonstration of how to do Pause, Prompt, Praise.
Supervised role-playing by parents using the Pause, Prompt, Praise method.

Parents were also given a summary of the method for reference. The teachers similarly made sure that children took home grade 1 books of their choice each day.

(d) *Direct Instruction (DI)*

The DI group met for three evening training sessions of one and a half hours each. These sessions included:

The provision of a Direct Instruction reading programme — *Teach Your Child to Read in 100 Easy Lessons* (TYCTOR) (Englemann *et al.*, 1983) — which is based on the fast-cycle component of DISTAR Reading I and II programmes (Englemann and Bruner, 1975) and written especially for parents.
An introduction to the basic concepts of Direct Instruction, the organisation of the programme and the lesson format.
A demonstration of some lessons from TYCTOR with an emphasis on correct letter pronunciation and blending, error correction procedures.

Supervised role-playing of selected lessons from TYCTOR by parents.

No reading material was brought home from school by the children in this group. The single book which constitutes TYCTOR was the only resource used.

The parents in all groups used social reinforcement for child effort and correct reading each day throughout the intervention period.

Measures

As the children were beginning readers, it was necessary to ascertain whether or not the groups of children were equivalent in their readiness to benefit from their parents' help. Pre-reading skills were, therefore, assessed. Phonic knowledge was tested by having the children complete List A from the Graded Spelling Test (Daniels and Diack, 1977) and a measure of reading readiness, which included verbal reasoning, word fluency, visual and auditory discrimination and word meanings, was taken from exercises they were given to complete from The Red Book (Thurstone, 1973). Reading accuracy and comprehension before and after the 10-week experimental period was assessed on the Neale Analysis of Reading Ability (1988). All testing was carried out by the second author.

RESULTS

The results of the pre-reading tests of phonic knowledge and reading readiness are presented in Table 1. One-way analysis of variance showed that the groups of children were not significantly different in their phonic skills at the beginning of the study [F(3,36) = 0.108, NS] nor in their readiness to read [F(3,36) = 0.017, NS]. Thus, there seemed to be no bias in the likelihood of improvement in favour of any one group.

TABLE 1

MEANS (AND STANDARD DEVIATIONS) OF STUDENTS' PRE-READING SKILLS BEFORE DIFFERENTIAL PARENT INVOLVEMENT WITH READING

Parent Reading Group	N	Phonic Knowledge	Reading Readiness
Hearing Reading	10	1.6 (1.56)	18.7 (1.42)
Paired Reading	10	1.6 (1.57)	18.7 (1.13)
Pause, Prompt, Praise	10	1.9 (1.64)	18.6 (1.20)
Direct Instruction	10	1.5 (1.50)	18.6 (1.28)

The pre- and post-test results on the Neale Analysis of Reading Ability are shown in Table 2. A one-way analysis of variance confirmed that the pretest scores for reading accuracy and comprehension by the groups were not significantly different [F(3,36) = 0.706, NS]. However, a subsequent analysis of covariance on the post-test scores, with pretest scores as the covariate, showed that the observed differences between groups at this stage were statistically significant (Accuracy [F(1,35) = 11.751, P<0.05]; comprehension [F(1,35) = 31.396, P<0.05]).

A number of planned comparisons between treatment groups were made according to the stated hypotheses (see Table 3). Applying the Modified Bonferroni Test (Keppel, 1982), the critical value of t was set at 2.101 (P<0.03) for all comparisons to correct for possible type 1 errors.

According to the diaries kept by the parents, the mean time spent with their children when reading ranged from 10.89 to 11.50 hours during the intervention period. There were no significant differences between groups on this measure [F(3,36) = 0.133, NS].

DISCUSSION

The results strongly suggest that increases in rates of reading progress can be expected if parents are taught more precise instructional methods that go beyond the provision of increased opportunities to practise, enhanced interest and reinforcement. In particular, there was demonstra-

TABLE 2

Means (and Standard Deviations) of Pre- and Post-Intervention Reading Ages (in Years) and Gains Made (in Months) by Group

Group	Pre-Test	Post-Test	Gain
	Reading Accuracy Age		
Direct Instruction	5.70 (0.03)	7.11 (0.47)	+16.8 (5.4)
Paired Reading	5.71 (0.04)	6.76 (0.36)	+12.6 (4.3)
Pause, Prompt, Praise	5.75 (0.03)	6.53 (0.53)	+ 9.9 (6.4)
Hearing Reading	5.70 (0.00)	6.19 (0.34)	+ 5.9 (4.1)
	Reading Comprehension Age		
Direct Instruction	5.70 (0.00)	7.06 (0.47)	+16.3 (5.6)
Paired Reading	5.70 (0.00)	6.84 (0.25)	+13.7 (3.0)
Pause, Prompt, Praise	5.70 (0.00)	6.52 (0.48)	+ 9.8 (5.8)
Hearing Reading	5.70 (0.00)	6.25 (0.39)	+ 6.3 (4.7)

TABLE 3

Comparisons of Post-test Mean Reading Accuracy and Reading Comprehension Scores Between Groups

Reading Accuracy	t-value	df	P
DI ⌓ PR	1.772	18	0.045 (NS)
DI > PPP	2.468	18	0.011
DI > HR	4.763	18	< 0.001
PR > HR	3.453	18	0.002
PPP ⌓ HR	1.631	18	0.058 (NS)
Reading Comprehension			
DI ⌓ PR	1.237	18	0.115 (NS)
DI > PPP	2.401	18	0.013
DI > HR	3.958	18	< 0.001
PR > HR	3.836	18	< 0.001
PPP ⌓ HR	1.314	18	0.102 (NS)

tion of the feasibility and value of brief training in Direct Instruction and Paired Reading techniques. For the expenditure of two hours of extra professional training time per family, beginning readers taught by these methods achieved rates of progress in reading that were two to three times greater than rates achieved by the HR group. These increases were not due to differences in children's "engaged time" on reading between groups and seem unlikely to have been due to the slight variations in time the parents spent on training.

Support for the Pause, Prompt, Praise method as an extra tutoring strategy was not strong. Although the group's mean gain was greater than the HR group, the results were not consistent enough across students to reach statistical significance. Clearly, more comparative studies of this kind need to be carried out before the value of teaching Pause, Prompt, Praise to parents of beginning readers over and above the minimal Hearing Reading strategy can be assumed.

At a theoretical level, the results indicate that the strategy of reading simultaneously with children, as in Paired Reading, has utility for accelerating early reading attainment. However, why and how Paired Reading works has not yet been determined. For example, is adult modelling of fluent reading most important, or is the immediate, corrective feedback given to the child when errors occur more critical? The contribution of these and other elements of the Paired Reading method are well worth identifying experimentally.

The complex set of programming strategies and instructional techniques embodied in Direct Instruction reading packages has received broad experimental support already (e.g., Carnine and Silbert, 1979), not to mention the extensive literature indicating the importance of phonological recoding strategies for beginning readers. Suffice it to say that the particular programme of "Teach Your Child to Read . . .", which was evaluated here for the first time, produced equally impressive results with regular students, when delivered by parents, as those obtained by other Direct Instruction programmes designed on the same principles which have been applied to disadvantaged children (Becker, 1977) and children with a range of learning difficulties (e.g. Lockery and Maggs, 1982) by professional teachers. Moreover, the gains made by the children were, on a test of general reading skills, unrelated to the programme content and were achieved despite the fact that the group completed only one half to three-quarters of the programme in the time available. This meant that they had been exposed to short reading passages, no longer than a few sentences in length, which were of very restricted, phonically controlled prose and written in a modified script at this early stage of the programme. Thus, the group had read far fewer regularly scripted passages at home than any of the other groups during the experimental period and had read no extra books as such. It is possible that the group would have gone much further ahead had the full TYCTOR programme been completed.

The issue of how to respond to children's reading errors needs proper experimentation. Nevertheless, the uniqueness of the PPP method is its dependence on prompting children's self-instructional responses to the errors they make. The equivalent progress of the PPP group to the HR group, particularly in reading accuracy, may have been due to the inefficiency of the method with beginning readers who had little experience of written prose and little or no phonic knowledge to draw on for self-correction. In view of the marked superiority of the DI and PR groups in reading accuracy and comprehension, however, real differences would seem to be due to the additional instructional components of the DI and PR methodologies, including their similar correction procedures. All that can be concluded, other than the fact that the PPP group did not progress as quickly as the DI group and the PR group, is that the self-correction strategies taught in the PPP methodology were not sufficient to increase the rates of gain of these beginning readers significantly beyond those achieved by a simpler Hearing Reading strategy that did not employ them. Perhaps, however, as with Direct Instruction, a better assessment of the utility of Pause, Prompt, Praise for parents would have been over a longer period of six months to one year, as this would have given the children more time to master the self-correction strategies and gain more reading experience from which to draw sufficient knowledge to apply them.

The generalisability of these results must be tempered by the small-scale nature of this study and the relatively short time period involved in the intervention. However, the significance levels of the inter-group differences do give some confidence in concluding that there is real value in teaching parents the PR strategy and the DI early reading package. Long-term follow-up measures are now required to assess the maintenance of these gains. Also, comparative studies need to be carried out with students who are at later stages of learning to read.

Correspondence and requests for reprints should be addressed to Mr David Leach, Psychology Department, School of Social Sciences, Murdoch University, Murdoch, Perth, Western Australia 6150.

REFRENCES

Becker, W. C. (1977). Teaching reading and language to the disadvantaged. *Harvard educ. Rev.*, 47, 518-545.
Becker, W. C., and Gersten, R. (1982). A follow-up of Follow-Through. *Am. educ. Res. J.*, 19, 75-92.
Becker, W. C., Engelmann, S., Carnine, D. W., and Maggs, A. (1980). Direct instruction technology: recent developments and research findings. In Karoly, P., and Steffens, J. (Eds.), *Advances in Child Behaviour and Therapy (Volume 2).* New York: Gardener Press.
Branwhite, A. B. (1983). Boosting reading skills by direct instruction. *Br. J. educ. Psychol.*, 53, 291-298.
Bush, A. (1983). Can pupils' reading be improved by involving their parents? *Remed. Educ.*, 18, 168-170.
Carnine, D. W., and Silbert, J. (1979). *Direct Instruction Reading*. Columbus, Ohio: Charles E. Merrill.
Clay, M. M. (1979). *Reading: The Patterning of Complex Behaviour*. (2nd ed.) Auckland, N.Z.: Heinemann.
Daniels, J. C., and Diack, H. (1977). *The Standard Reading Tests*. London: Hart-Davis.
Englemann, S. (1980). *Direct Instruction.* Englewood Cliffs, N.J.: Prentice-Hall.

Englemann, S. and Bruner, E. (1975). *DISTAR Reading I & II: An Instructional System.* Chicago: Science Research Associates.

Englemann, S., Haddox, P., and Bruner, E. (1983). *Teach Your Child to Read in 100 Easy Lessons.* New York: Simon and Schuster.

Francis, H. (1987). Hearing beginning readers read: problems of relating practice to theory in interpretation and evaluation. *Br. educ. Res. J.*, 13, 215-225.

Gregory, P., Hackney, C., and Gregory, N. M. (1982). Corrective reading programme: an evaluation. *Br. J. educ. Psychol.*, 52, 33-50.

Hannon, P. (1987). A study of the effects of parental involvement in the teaching of reading on children's reading test performance. *Br. J. educ. Psychol.*, 57, 56-72.

Hewison, J., and Tizard, J. (1980). Parental involvement and reading attainment. *Br. J. educ. Psychol.*, 50, 209-215.

Keppel, G. (1982). *Design Analysis: A Researcher's Handbook.* Englewood Cliffs, N.J.: Prentice-Hall.

Lockery, M., and Maggs, A. (1982). Direct instruction research in Australia: a ten year analysis. *Educ. Psychol.*, 2, 263-288.

Love, J. M., and Van Biervliet, A. (1984). Training parents to be home reading tutors: generalization of children's reading skills from home to school. *Except. Child*, 31, 114-127.

McNaughton, S. (1988). A history of errors in the analysis of oral reading behaviour. *Educ. Psychol.*, 8, 21-30.

McNaughton, S., Glynn, T., and Robinson, V. M. (1981). *Parents as Remedial Reading Tutors: Issues for Home and School.* Wellington, N.Z.: New Zealand Council for Education Research.

Morgan, R., and Lyon, G. (1979). Paired reading: a preliminary report on a technique for parental tuition of reading with retarded readers. *J. Child Psychol. Psychiat.*, 20, 151-160.

Neale, M. D. (1988). *Neale Analysis of Reading Ability.* Melbourne: Australian Council for Educational Research.

Noon, L., and Maggs, A. (1980). Accelerating written language processes in normal and gifted children: direct instruction strategies and sequences. *Reading Educ.*, 5, 11-19.

Quisenberry, N. L., Blakemore, C., and Warren, C. A. (1977). Involving parents in reading: an annotated bibliography. *Reading Teach.*, 31, 34-39.

Robson, D., Miller, A., and Bushell, R. (1984). The development of paired reading in High Peak and West Derbyshire. *Remed. Educ.*, 19, 177-183.

Share, D. L., Jorm, A. F., Matthews, R., and MacLean, R. (1987). Parental involvement in reading progress. *Aust. Psychol.*, 22, 43-51.

Thurstone, T. G. (1973). *The Red Book.* Sydney: Science Research Associates.

Tizard, J., Schofield, W. N., and Hewison, J. (1982). Collaboration between teachers and parents in assisting children's reading. *Br. J. educ. Psychol.*, 52, 1-5.

Topping, K., and Wolfendale, S. (Eds.) (1985). *Parental Involvement in Children's Reading.* London: Croom Helm.

Wheldall, K., and Glynn, T. (1988). Contingencies in contexts: a behavioural interactionist perspective in education. *Educ. Psychol.*, 18, 5-19.

(*Manuscript received 16th August*, 1988)

Br. J. educ. Psychol., 60, 356-363, 1990

CULTURAL SPECIFICITY OF APPROACHES TO STUDY

By DAVID KEMBER and LYN GOW
(Hong Kong Polytechnic, Hong Kong)

SUMMARY. This paper discusses the cultural specificity of constructs reported in research into approaches to study of western students. Data are reported from a survey of approaches to study of students at a Hong Kong tertiary institution. The Biggs' Study Process Questionnaire (SPQ) was administered to a sample of 1043 students. The SPQ and the Approaches to Studying Inventory were both administered to a smaller sample to aid in interpretation of the data. The resulting factor structures for deep and achieving approach scales were reasonably consistent with those obtained in western countries. However, the results cast doubt on the direct transferability of the surface approach construct to Hong Kong students for whom a narrow approach appeared to predominate, characterised by a systematic, step-by-step, processing of information. Students tried to understand each segment of information before committing it to memory. It is suggested that the different approaches might arise because of either limited English language ability or the nature of schooling and upbringing of children in Hong Kong.

INTRODUCTION

Various instruments have been developed to examine student approaches to learning. The two which seem to have been most extensively used for higher education are the Approaches to Studying Inventory (ASI) (Ramsden and Entwistle, 1981) and the Study Process Questionnaire (SPQ) (Biggs, 1987). Both these questionnaires have been administered to substantial numbers of students in Australia and the United Kingdom (e.g., Entwistle and Ramsden, 1983; O'Neil and Child, 1984; Harper and Kember, 1986; Biggs, 1987). The instruments, however, have not been used extensively outside these countries. Phenomenographic studies (Marton, 1981) of student approaches to learning have also concentrated on students in western countries, principally Australia, the United Kingdom and the Scandinavian countries (e.g. Marton *et al.*, 1984).

Therefore, the constructs which have emerged from the extensive research into student approaches to learning can confidently be assumed to be applicable to students in western countries. There can be less confidence, however, about their applicability to other areas of the world where students are influenced by different cultural traditions, different educational systems, and where the students might not use their first language for study.

The small amount of literature which does relate to non-western countries does not give a clear verdict on the universality of the instruments and associated constructs. Hattie and Watkins (1981) investigated the internal structure of the SPQ with Australian and Filipino students. The Australian data yielded two-, three- and six-factor solutions, interpretable according to Biggs' model. For the Filipino sample, only a two-factor solution was interpretable and the scales had low to moderate internal consistencies, which led Hattie and Watkins to question the appropriateness of the SPQ for Filipino students. Watkins *et al.* (1986) used a short version of the ASI with Filipino school students and used confirmatory factor analysis to test the fit between factor structure and that which would be expected from Entwistle's model. They concluded that a good fit had been demonstrated. Entwistle and Kozéki (1985) used a shortened version of the ASI with school students in Britain and Hungary. They reported that the factor structure of the inventory was remarkably stable across nationality. Meyer and Parsons (1989) used the ASI at the Cape Technicon in South Africa, but recognised that their population "shares many of the linguistic and cultural characteristics" (p. 151) of Entwistle and Ramsden's original population. Their second order factor structure of the ASI subscales was similar to those found by other researchers.

This paper reports an investigation of the approaches to study of students at a tertiary institution in Hong Kong. The project originally used the SPQ in an attempt to characterise study approaches of an extensive sample. The ASI and the SPQ were subsequently administered to a smaller sample in an effort to clarify the meaning of results from the SPQ.

METHOD

Questionnaires

The 42 items of the SPQ were translated into Chinese for the survey by a research team at the Hong Kong University, led by Professor John Biggs. Contributions to the translation were made by a number of people, before independent back translation. The ASI was translated by a team of four, followed by independent back translation and further modification. Both questionnaires given to students had the item statements in English, followed immediately by the Chinese translation.

Responses to the items of both questionnaires were made by ticking responses to a five-point scale on a separate cover sheet. The reverse of the cover sheet contained 18 questions requesting personal details, grades in two English language examinations and self-ratings in reading, writing and speaking English. The students' name, student number and Hong Kong identity card number were requested so that the data from the questionnaires could be compared with examination grades. In addition the research project is longitudinal in nature so there is an intention to re-administer the questionnaire.

Sample

The sample for the SPQ consisted of first- and final-year students in degree level courses at a tertiary institution in Hong Kong. The students were from the departments of Accountancy. Applied Social Studies, Diagnostic Sciences, Language and Communication, Rehabilitation Sciences, and Textiles and Clothing. The number of questionnaires with sufficient responses to calculate SPQ scores for each of the six subscales of the SPQ was 1043. The total enrolment in the selected classes in the chosen departments was 1294. The usable response rate for the survey was 80.6 per cent.

A second group of 159 students from the departments of Accountancy, Diagnostic Sciences and Language and Communication completed both the ASI and the SPQ. The return for this sample was 84.0 per cent.

Administration

The SPQ was administered during the third or fourth weeks of the academic year. The joint administration of the ASI and SPQ took place close to the start of the second semester. In both instances, lecturers handed the questionnaires and cover sheet to students in lectures. Students were asked to complete the questionnaires before leaving the lecture.

Semi-structured interviews

In addition to the questionnaires, data were gathered by semi-structured interviews with 20 randomly selected students from the sample which completed the SPQ. Questions in the interview schedule probed the students' motives and strategies for study, such as approaches to reading and note taking. Within these broad areas students were free to raise any issues which were of concern to them. Any prompting was aimed at obtaining more information on points raised by students, and was kept as non-directive as possible. Interviews were conducted in Cantonese so that the students could converse freely. The interviews were tape-recorded. Transcripts in English were prepared from each tape by the bilingual interviewer.

Participants' responses were then sorted into broad headings. More detailed sorting of comments into common concerns and constructs followed. The sorting was conducted by three independent judges. The constructs discovered by the judges were then compared for consistency, and a process of peer consultation (Owens, 1982), within the research group, was used to evaluate the findings. The principles for sorting and coding interview data are discussed by Grove (1988). Constructs which emerged from the interview data were then cross-checked against results from the questionnaire data as a triangulation procedure (Guba, 1981).

RESULTS

Questionnaire data were analysed with SPSS-X (SPSS Inc., 1986). The SPQ yields scores for each student on six subscales by summing seven items in each case. The ASI has 16 subscales (Ramsden and Entwistle, 1981).

Reliability of SPQ scales and subscales

The reliability, in terms of internal consistency, of the SPQ scales and subscales was determined using Cronbach's α coefficient. The results are reported in Table 1.

TABLE 1
RELIABILITY OF SPQ SCALES AND SUB-SCALES

SPQ Scale	
Surface	0.72
Deep	0.80
Achieving	0.79
SPQ Sub-scale	
Surface motive	0.61
Surface strategy	0.57
Deep motive	0.66
Deep strategy	0.72
Achieving motive	0.73
Achieving strategy	0.76

Factor analysis of SPQ

For the SPQ the methods of Biggs and O'Neil and Child (both reported in O'Neil and Child, 1984) were replicated in the present study, so that comparison could be made with previous factor analyses. The SPSS-X package (SPSS Inc., 1986) was used to perform principal factor analysis, with iterations to obtain an estimate of communalities, followed by varimax orthogonal rotation for factors with eigenvalues greater than one. The 10-factor solution is shown in Table 2. For purposes of comparison, the table also shows results from Biggs' Australian university and college of advanced education (CAE) samples and O'Neil and Child's British polytechnic sample (O'Neil and Child, 1984). O'Neil and Child obtained a 13-factor solution but only report the six factors they were able to interpret.

Factor analysis of the ASI

A principal component analysis with oblique rotation was carried out on the 16 subscales of the ASI. Five factors were extracted, and only items with loadings above 0.30 are reported in Table 3.

Table 3 also shows the factor structure obtained when the ASI was administered to a sample of senior secondary school students in Hong Kong (Lai, 1989).

DISCUSSION

Factor analysis of the SPQ

The factor structure for the SPQ obtained in the present study clearly differs from those previously reported for Biggs' study and for O'Neil and Child's (1984) British sample. Factors 1, 2 and 3 correspond reasonably well with the subscales achieving strategy, achieving motive and deep strategy respectively. Factors 4 to 9, however, all contain items from two or three subscales.

O'Neil and Child expressed some doubts about the surface dimension based on its lower α coefficient and the factor structure they obtained. In the present study, the surface subscales are not discrete: surface strategy items appearing with surface motive and/or achieving motive. Factors 4 and 8 contain items from both deep motive and strategy subscales.

Factor analysis of the ASI

Factor structures from the ASI have recently been reviewed by Harper and Kember (1989) and Meyer and Parsons (1989). Neither paper supports the original interpretation of Entwistle and Ramsden (1983).

Factor 2 of this study corresponds with the "deep orientation" of Harper and Kember's (1989) review or the "meaning orientation" of Meyer and Parsons (1989). The "surface orientation" of Harper and Kember or the "reproducing orientation" of Meyer and Parsons do not appear in this study despite appearing consistently in other studies covered by the two reviews (Morgan *et al.*, 1980; Ramsden and Entwistle, 1981; Watkins, 1982, 1983; Watkins and Hattie, 1985; Clarke,

TABLE 2

FACTOR ANALYSIS OF THE SPQ ITEMS

Sub-scale	Question no.	Biggs' (university)						Biggs' (CAEs)						O'Neil and Child						Hong Kong Polytechnic									
		1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	7	8	9	10
SURFACE MOTIVE	1	655						671						496										307					
	7																									479			
	13	608				347		607				290		437				337						317					
	19																									621			
	25		594							-537															432				
	31	392		-476						-477				508															
	37	687						686						559										710				340	
SURFACE STRATEGY	4		678						369						606										495				
	10																									410			
	16																												
	22		557				329								567										600				
	28		368						514					342														313	
	34								723					361	341		-476											347	
	40								714					401														494	
DEEP MOTIVE	2			378						398	232					434				289			502						
	8									304	360					296	397										318		
	14			337						447						480							382						
	20			397						390						625							547						
	26			556	334					398						549							364						
	32															276	372										378		
	38	272																									443		
DEEP STRATEGY	5				428						650						477					304					379		
	11				603					332	318											544							
	17			246	261		383			268	385		325				404						316				272		
	23				728						708						480					616							
	29		-304				455			347			515							255			443						
	35		-377				335						520			310						283							476
	41				701						650						473					622							
ACHIEVING MOTIVE	3					734						707						435			361			298					
	9					649	347					618						481	415		364								
	15					643						600						385			676								
	21									306		321						376											
	27											494																	
	33					438						705						572			621							366	
	39											398						327						256				445	
ACHIEVING STRATEGY	6						340						573							393									
	12						606						474						683	560									
	18						676						539						625	613									
	24						680						677						536	689									
	30				315		506						539						278	377		296							
	36						649						704							432									
	42						572						447						587	479									
% variance explained																				14.2	7.8	3.2	2.3	1.9	1.4	1.3	1.2	1.1	0.8

Decimal points and loadings <0.25 omitted

TABLE 3

FACTOR ANALYSIS OF THE ASI SUB-SCALES

	This project					Hong Kong secondary school students (Lai, 1989)				
	1	2	3	4	5	1	2	3	4	5
Deep approach		65				67				
Relating ideas		67				67				
Use of evidence		71				65				
Intrinsic motivation		70				61				
Surface approach					–69		48			
Syllabus-boundness	66									–46
Fear of failure	40				–30		55			
Extrinsic motivation		–34	65						43	
Strategic approach			48	36	–52				37	
Disorganised study methods				72				45		
Negative attitude			31		73					
Achievement motivation			77					61	50	
Comprehension learning		44		61				33		57
Globetrotting				30	–55			36		
Operation learning	76						35			–32
Improvidence	74						63			

Decimal points are not shown
Only factor loadings greater then 0.30 are shown.

1986; Harper and Kember, 1986; Meyer, 1988; Meyer and Parsons, 1989). Instead a factor, labelled "narrow orientation" by Harper and Kember (1989), which has usually appeared third was the first factor to emerge. The factor has the heaviest loadings on operation learning and improvidence and also loads on syllabus boundness and fear of failure.

The subscale operation learning is derived from the work of Pask (1976a) and refers to an ordered systematic approach to study. The term improvidence is used to describe the corresponding learning pathology, symptoms of which are failure to build an overall picture or concept map of the elements of a subject, and the absence of any vision of the inter-relationship between topics. Pask sees improvident learners as failing to make use of appropriate analogies.

The third factor is close to the "goal orientation" factor defined by Harper and Kember as loading primarily on the extrinsic and achieving motivation subscales. The fourth factor of the present study has the heaviest loading on disorganised study and also loads on comprehensive learning and its corresponding pathology of globe-trotting. The factor appears to describe students opposite in type to factor one. The students have a holistic disorganised approach. The fifth factor contains items normally found in the "surface orientation" factor, but with improvidence and syllabus boundness removed.

Comparison with previous factor structures

For both the SPQ and ASI there is a reasonable correspondence between the present study and previous studies on the factors relating to a deep approach. The consistency between the studies, in different cultural contexts, possibly results from consistency in the goals of higher education. Gow and Kember (1990) drew on a number of sources (Entwistle and Percy, 1971, 1974; Percy and Salter, 1976; Australian Government Department of Employment, Education and Training, 1987; Baird, 1988) to show that qualities loosely embodied in the term "independent learning" are widely recognised as characterising the goals of higher education.

The documented goals of tertiary courses in Hong Kong are remarkably similar to those elsewhere (Gow *et al.*, 1989). The deep orientation factor does not correspond precisely with the qualities of independent learning, but it is at least a pre-requisite for independence. The ideal student, judged by faculty members, might therefore be seen as having similar characteristics, as measured by deep approach scales, regardless of the cultural setting.

The situation is quite different, however, with the less desirable study traits. The ASI is easier to interpret in this instance. For this sample the "narrow orientation" factor appears to assume greater significance than the "surface orientation" factor which has appeared consistently in first or second position in other studies. To understand the reason it is necessary to first amplify the meaning associated with the factor.

The semi-structured student interviews aid the interpretation. A typical comment came from an accountancy student:

> I read in detail section by section. If I find any difficulties I try my best to solve the problem before I go onto the next section. . . . If you don't memorise important ideas when you come across them then you will be stuck when you go on. You must memorise and then go on — understand, memorise and then go on — understand, memorise and then go on. That is my way of studying.

The approach is related to surface learning, since it relies on the memorisation of content defined by the lecturer. However, surface learners do not necessarily seek to understand what they are learning by rote. Indeed an element of Biggs' (1987) characterisation of a surface learner is one who "avoids personal or other meanings the task may have" (p. 15). The accountancy student quoted above clearly does not fit this characterisation. The distinguishing features of the narrow orientation are therefore:

* learning tasks are defined by the lecturer
* understanding is sought in a narrow but systematic, step-by-step approach
* as each part of the task is understood, an attempt is made to memorise the pertinent details.

The existence of the narrow orientation, related to but distinct from a surface approach can be advanced as a reason for the splitting of the items in the surface approach scales in the SPQ. O'Neil and Child (1984) expressed some reservations about the surface approach scale following use with students in western countries. With the emergence of the narrow orientation in the present study, the surface approach factors presumably split into a number of related constructs. As the SPQ does not discretely draw upon Pask's (1976a, 1976b) styles and pathologies, the relationship to the narrow orientation is less apparent.

Possible reasons for cultural differences

There are two possible explanations for the strengthening of the narrow orientation at the expense of the surface orientation. The first concerns the English language ability of the students. The first language of the sample is almost invariably Cantonese (Gibbons, 1987). English is very much restricted to formal interaction within the classroom. As the population of Hong Kong are almost all Cantonese speakers, English is used so little that few students have acquired the level of fluency in the language which qualifies it as a second language. Rather, the limited usage of English in Hong Kong in general means it is effectively an auxiliary language (Luke and Richard, 1981) rather than a second language.

An English language scale (Kember *et al.*, 1989) was used as a measure of the students' English language ability in a related study, to be reported in more detail elsewhere. The English language scale was found to be correlated significantly with subscales in the ASI shown in Table 4.

TABLE 4
CORRELATION BETWEEN ASI SUB-SCALES AND THE ENGLISH LANGUAGE SCALE

ASI sub-scale	r	P<
Improvidence	−0.32	0.001
Syllabus – boundness	−0.29	0.001
Negative attitude	0.25	0.5
Fear of failure	−0.24	0.05
Surface approach	−0.23	0.05

The English language scale has significant negative correlations with three of the four ASI subscales in the narrow orientation scale. It is possible that those with a poor command of English find it difficult to scan through a document to find the main points or, in a lecture, are unable to sense the key concepts. Instead the students might work step-by-step through a reading concentrating initially on deciphering the rhetorical aspects (i.e., letters, words and sentences) of the text. In a lecture the students often concentrate on copying from a blackboard or overhead projection (Gow and Kember, 1990). The students' approach to knowledge acquisition may therefore be quite blinkered. All information has to be read in a systematic linear manner, with an attempt made to comprehend each constituent part.

The second possible explanation is that, as a result of schooling and/or cultural tradition, the Hong Kong students have a high regard for authority and are therefore comfortable with a regulated approach to study (Ho, 1986; Murphy, 1987; Dunbar, 1988; Tobin *et al.*, 1989).

The findings of this project suggest that the deep or meaning orientation scales of both instruments can be used with some confidence in settings other than western countries, as the underlying constructs seem to be stable across educational systems. There do appear to be some cultural influences on the less desirable study approaches. The SPQ, and particularly the ASI, have provided some insights into the meaning of these other orientations. There would seem to be no reason to discourage the use of these instruments in non-western countries, provided care is taken with interpretation. There is obviously a need though for further work if the surface-related dimensions of the two instruments are to be unambiguously interpreted when applied to students in these parts of the world.

Acknowledgements.—This research was carried out with a grant from the Hong Kong Polytechnic. Rosalia Chow, Irene Siaw and Roger Liu assisted with data collection and analysis. The research team would like to thank Professor John Biggs for supplying the SPQ questionnaire for the project. The team acknowledges the fruitful collaboration with the research team at the City Polytechnic of Hong Kong who are undertaking similar research into the approach to study of their students.

Correspondence and requests for reprints should be addressed to Dr David Kember, Hong Kong Polytechnic, Hung Hom, Kowloon, Hong Kong.

REFERENCES

Australian Government Department of Employment, Education and Training (1987). Policy discussion paper on higher education. Canberra: Australian Government Printing Service.

Baird, J. R. (1988). Quality: what should make higher education "higher"? *Higher Educ. Res. Dev.*, 7 (2), 141-152.

Biggs, J. (1987). *Student Approaches to Learning and Studying*. Melbourne: Australian Council for Educational Research.

Clarke, R. M. (1986). Students' approaches to learning in an innovative medical school: a cross-sectional study. *Br. J. educ. Psychol.*, 56, 309-321.

Dunbar, R. (1988). Culture-based learning problems of Asian students: some implications for Australian distance educators. *ASPESA Papers*, 5, 10-21.

Entwistle, N. J., and Kozéki, B. (1985). Relationships between school motivation, approaches to studying, and attainment, among British and Hungarian adolescents. *Br. J. educ. Psychol.*, 55, 124-137.

Entwistle, N. J., and Percy, K. A. (1971). Educational objectives and student performance within the binary system. In Page, C. F., and Gibson, J. (Eds.), *Research into Higher Education 1970*. London: SRHE.

Entwistle, N. J., and Percy, K. A. (1974). Critical thinking or conformity? An investigation of the aims and objectives of higher education. In Page, C. F., and Gibson, J. (Eds.), *Research into Higher Education 1973*. London: SRHE.

Entwistle, N. J., and Ramsden, P. (1983). *Understanding Student Learning*. London: Croom Helm.

Gibbons, J. (1987). *Code-mixing and Code Choice: a Hong Kong Case Study*. Clevedon, Avon: Multilingual Matters Ltd.

Gow, L., and Kember, D. (1990). Does higher education promote independent learning? *Higher Educ.*, 19, 307-322.

Gow, L., Balla, L., Kember, D., Stokes, M., Stafford, K., Chow, R., and Hu, S. (1989). Approaches to study of tertiary students in Hong Kong. *Bull Hong Kong Psych. Soc.*, 22 (23), 57-77.

Grove, R. W. (1988). An analysis of the constant comparative method. *Qualitative Studies in Education*, 1 (3), 273-279.

Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic enquiries. *Educ. Commun. Technol. J.*, 29 (2), 75-91.

Harper, G., and Kember, D. (1986). Approaches to study of distance education students. *Br. J. educ. Technol.*, 3 (17), 212-222.

Harper, G., and Kember, D. (1989). Interpretation of factor analyses from the approaches to studying inventory. *Br. J. educ. Psychol.*, 59, 66-74.

Hattie, J., and Watkins, D. (1981). Australian and Filipino investigations of the internal structure of Biggs' new Study Process Questionnaire. *Br. J. educ. Psychol.*, 51, 241-244.

Ho, D. Y. F. (1986). Chinese patterns of socialisation: a critical review. In Bond, M. H. (Ed.), *The Psychology of the Chinese People*. London: Oxford University Press.

Kember, D., Gow, L., Chow, R., Siaw, I., Barnes, P., and Hunt, J. (1989). Approaches to study of students whose first language is not English: preliminary findings. In Bickley, V. (Ed.), *Teaching and Learning Styles Within and Across Cultures: Implications for Language Pedagogy*. Hong Kong: Institute for Language in Education.

Lai, T. (1989). Effects of study strategies, orientating aids and concept maps as advance graphic organisers on learning from text. Unpublished Masters degree thesis, Chinese University of Hong Kong.

Luke, K. K., and Richards, J. C. (1981). English in Hong Kong: functions and status. *English Worldwide*, 3 (1), 46-64.

Marton, F. (1981). Phenomenography-describing conceptions of the world around us. *Instruct. Sci.*, 10, 177-200.

Marton, F., Hounsell, D., and Entwistle, N. (1984). *The Experience of Learning*. Edinburgh: Scottish Academic Press.

Meyer, J. H. F. (1988). Student perceptions of learning context and approaches to studying. *South African J. Higher Educ.*, 2, 73-82.

Meyer, J. H. F., and Parsons, P. (1989). Approaches to studying and course perceptions using the Lancaster Inventory — a comparative study. *Studies in Higher Educ.*, 14 (2), 137-153.

Morgan, A., Gibbs, G., and Taylor, E. (1980). *The Work of the Study Methods Group*. (Study Methods Group Report No. 1). Milton Keynes: The Open University, Institute of Educational Technology.

Murphy, D. (1987). Offshore education: a Hong Kong perspective. *Aust. Univ. Rev.*, 30 (2), 43-44.

O'Neil, M. J., and Child, D. (1984). Biggs' SPQ: a British study of its internal structure. *Br. J. educ. Psychol.*, 54, 228-234.

Owens, R. (1982). Methodological rigor in naturalistic inquiry: some issues and answers. *Educ. Admin. Q.*, 18 (2), 1-21.

Pask, G. (1976a). Styles and strategies of learning. *Br. J. educ. Psychol.*, 46, 128-148.

Pask, G. (1976b). Conversational techniques in the study and practice of education. *Br. J. educ. Psychol.*, 46, 12-25.

Percy, K. A., and Salter, F. W. (1976). Student and staff perceptions and "the pursuit of excellence" in British higher education. *Higher Educ.*, 5, 457-473.

Ramsden, P., and Entwistle, N. J. (1981). Effects of academic departments on students' approaches to studying. *Br. J. educ. Psychol.*, 51, 368-383.

SPSS Inc. (1986). *SPSS-X Users Guide*. Chicago: SPSS Inc.

Thomas, P., and Bain, J. (1984). Contextual differences of learning approaches: the effects of assessments. *Hum. Learning*, 3, 227-240.

Tobins, J. J., Wu, D., and Davidson, D. (1989). *Preschool in Three Cultures: Japan, China and the United States*. New Haven, Conn.: Yale University Press.

Watkins, D. (1982). Identifying the study process dimensions of Australian university students. *Aus. J. Educ.*, 26 (1), 76-85.

Watkins, D. (1983). Assessing tertiary study processes. *Hum. Learning*, 2, 29-37.

Watkins, D., and Hattie, J. (1985). A longitudinal study of the approaches to learning of Australian tertiary students. *Hum. Learning*, 4, 127-141.

Watkins, D., Hattie, J., and Astilla, E. (1986). Approaches to studying by Filipino students: a longitudinal investigation. *Br. J. educ. Psychol.*, 56, 357-362.

(*Manuscript received 15th January*, 1990)

BOOK REVIEWS

ANDERSON, E. S. (1989). *Speaking with Style: the Sociolinguistic Skills of Children.* London: Routledge, £30.00 hbk. ISBN 0-415-02256-8.

It will surprise few readers to learn that children aged from four to seven years are already able to use appropriate language when involved in social role play of a familiar sort. Thus, when playing mother in a game of "families" they can say things like: "Daddy's gonna come home really late. He has to be working," (age 5:0), and when playing a doctor: "I think we better give her an operation, all right?" (age 5:3). Not only do children use appropriate topics and words in such play, they also vary their intonation and other paralinguistic features and sometimes their syntax in appropriate ways. Sociolinguists such as Labov have previously argued that pre-adolescents are virtually "mono-stylistic" in their speech, thus presenting Elaine Slosberg Anderson with a neat opportunity to show how wrong they are.

This is a short book (175 pages) and admirably straightforward. Its own "register" is that of the academic empirical report. The author reports on her study of 18 white middle class Californian children in which she used what she calls "controlled improvisation" to study the development of register in their speech. This involved individual sessions of role play with hand-puppets, dressed to represent different social roles, including parents and child, doctor, nurse and patient and teacher and pupils. In each session the child was invited to take on the role of two of the three puppets, whilst the investigator played the third role. This allowed the children considerable freedom in what they did and said but also allowed the investigator to support the conversation and play with helpful and appropriate moves and comments.

The main outcomes are predictable: children are better at marking register in role when they are familiar with the role in question and as they get older: boys, at least in contemporary California, tended to avoid playing the roles of mother, nurse or patient, girls were prepared to try more roles and generally spoke more. The author argues that children generally, and boys in particular, preferred "high status" roles. One wonders if "status" is the right word, or whether it is simply power that attracts them, in this case the power to dominate the play action.

The more interesting findings are those which examine the detailed linguistic markers of social register that the children used. It seems that intonation and paralinguistic features in general lead the way. Children know from an early age what different people ought to sound like, in the sense of using a deeper and louder voice for adult males and high pitch, nasalisation and whining for a young child. At younger ages they tended to mark register chiefly when changing from one role to another; older children were more consistent.

Here, as throughout their social role play, what children produce is actually the social stereotype, rather than the reality that they have experienced. Thus children of working professional mothers still play "mother" as a feminine, nurturant stay-at-home figure, much given to terms of endearment such as "sweetie". The author feels that this is probably a developmental phenomenon related to early undifferentiated concepts of gender roles rather than an accurate reflection of culture. Besides intonation, the children sometimes varied lexical and syntactic features appropriately and the resulting conversations showed an intuitive grasp of who is likely to speak most, and to whom, in such social situations. The subjects were particularly good at "baby-talk" but found the appropriate register for a teacher or a foreign child much harder.

There is rather little theoretical discussion in this work and the investigation is not extended to more naturalistic situations. The reader is left to contemplate the fact that when children learn to talk they are learning the whole social "script" for familiar situations, rather than advancing one linguistic dimension at a time. Thus they know something of the paralinguistic, phonological, lexical, syntactic and discourse features of "scripts" such as family life or a visit to the doctor even at the age of four. On the other hand they often get details wrong which is why we, as adults, confidently expect their play to be amusing, as when the doctor, aged 4:2 says: "Where should we put her temperature? I forgot. Down her bra?" This is a book which will interest chiefly the specialist in language development. The price seems exorbitant.

RICHARD FOX

CASE, C., and DALLEY, T. (Eds.) (1990). *Working with Children in Art Therapy.* London: Routledge, £14.99 pbk. ISBN 0-415-01738-6

The ten contributors to this book are all experienced art therapists, each with much to offer in their own right. From the richness of the text, one is left with the feeling of almost too much experience compressed into these pages. It reads like a very stimulating conference.

The editors have attempted to place the practice of art therapy in context institutionally, culturally and to some extent theoretically.

The role of art therapy is discussed in relation to schools, the social services and the health service. Most chapters include summaries of relevant theoretical or historical background before going into examples. There are also specific references to issues of race and gender as they affect the child's world and the therapeutic relationship. While much of the book deals with individual therapy it also touches upon group therapy and family systems work. The child's family and institutional context is never far from awareness, even within individual treatment.

It is this view of art therapy as a contextualised practice which, for me, makes this an interesting and significant book. It offers the reader both a grasp of the state of play in art therapy and a feel for the realities of practice via its numerous case studies.

At the core of the work is the "safe place" of the therapeutic session. Within this are the pliable materials with which the child works — paper, water, clay, paint, glue — and the art therapist as a non-judgmental figure. The therapist stands by patiently as the child works, helping to hold together and give meaning to what emerges. The work opens a non-verbal channel of expression to the child who is able to expose fantasies and anxieties that may be unmanageable in everyday circumstances.

Interaction with materials is the common factor, but this may vary from highly articulate symbolic picture making, to a more elemental use of materials for their basic properties of wetness, sliminess, softness, etc.

Control and release, cleanliness and mess are important dimensions in the work. A disabled child with little sense of boundary, of self and non-self, may learn through therapy to distinguish say, edible from non-edible materials, whereas a frozen, over-controlled child may learn to express her inner sense of chaos by sliding around in slimy paint.

The phenomenon of messiness recurs in the book and is one example of the way in which cultural and institutional factors impinge. For instance, the emotional safety of the therapeutic session enables a girl to be messy, breaking a social taboo that puts pressure upon girls to be clean and tidy.

In a school environment the emotionally deprived child is liable to be very disruptive, yet the teacher has not the time or the means to address the anxieties underlying this behaviour. By withdrawing from the school environment for short periods each week, the child can expose and share the "bad bits" of him or herself without fear of repercussions.

Accounts in this book of the way in which visual metaphors are invented by the child and recognised by the therapist are often surprising and moving. The value of people working as therapists whose training is in visual art is clear in this respect, since art deals so strongly in visual metaphor and association. One is struck by the inventiveness not only of the children but of the therapists in their capacity to think flexibly about what is produced.

One of the values of imagery is that it can contain ambiguity — an ambiguity that matches the ambiguity of feelings. A girl depicts a cat threatening to devour a fish and can feel herself at different times to be both creatures, victim and aggressor.

Often the use of imagery is much more explicit than anything a child could say. When a black child depicts himself as a small grey blob in the corner of a piece of paper, he gives evidence of his emotional and cultural deprivation and his consequently diminished sense of self, which he could not articulate in words.

Another value of working with materials is that they exist outside the self and yet can be made to represent what is inside. The survival of the work allows the feelings themselves to survive and to be held, the therapist keeping the work from one session to another often being a crucial act of support.

But whereas many of the children described are able to explore their fantasies and feelings through symbolic activity in play and in visual representation, certain groups described are unable to do so. These range from children with acute disability and deprivation who find difficulty in the simplest task, say filling a bag with clay, to children with unusual drawing ability whose very skill is a measure of their deprivation, since they are able only to represent the world they see as surfaces. They have not developed the child's normal ability to symbolise, to play and to use the concept of "as if".

This book gives a convincing case for the role of art therapy within the education system and other services. The methods used could offer models for wider application, for instance in art education itself. The book will be essential reading for those in a position to employ art therapists, for students and for other professionals who work with children.

CHRIS CRICKMAY

DILLON, J. T. (1990). *The Practice of Questioning*. London: Routledge, £12.95 pbk. ISBN 0-415-04379-4.

This book invites the reader to "take a tour" of "a whole world of questioning where people think and do marvellous things with question and answer." The author is an authority on questioning and editor of a journal (*Questioning Exchange*) which devotes itself entirely to research in this field. The whole enterprise assumes, with a Californian optimism, that an encyclopaedic approach to questioning will necessarily yield interesting and useful information. It is the *Reader's Digest* approach to research, in which all enquiry is presumed to be valuable and all readers are expected to bring with them an omnivorous appetite for knowledge in all its shapes and forms, whether trivial or profound.

Since most psychologists and teachers see themselves as professionally involved in questioning it is reasonable to suppose that they will find at least some of the information in this book of interest. Yet it might have been a better book if either it had concentrated on addressing a single professional audience or it were condensed into about half its length, or both. As it is, the book addresses anyone and everyone, from journalists to lawyers, from personnel interviewers to detectives and tries to please them all. There are some gems in these 275 pages of text but their discovery requires much sifting.

Classroom questioning is the first "domain" to be sampled but it is one of the shortest chapters and provides one of the lowest yields for reader effort invested. The classroom is a complex place, we learn, and includes many contexts for different uses of questioning, hardly any of which involve a request for information. Teachers tend to dominate classroom discourse via question but it remains unclear whether or not teacher questions actually achieve their avowed purpose of helping pupils to learn. A student wishing to review this research might do better to consult another source, such as Pollard and Tann (1987), chapter 9. The main nuggets we carry away from this section are the following: teachers should think more about why they are asking questions and what they will do with the answers; they should spend more effort in listening to the answers they get; in order to get pupils to talk, in the sense of discussion, it is more effective *not* to ask questions but to use almost any other form of utterance, especially the statement. Interestingly enough, much the same conclusions turn out to be applicable in many other domains.

I enjoyed my brief coach-tour through the alternative professional worlds of the psychotherapist, advocate, policeman, journalist and so on. Each stop on the way provided some amusing anecdotes and that mixture of marvel and respect that comes from finding out both how little and how much makes up the knowledge and skill of any "professional". Trying to get people to say things that they don't want to say seems to be a depressingly common aim in many different occupations and is thoroughly difficult to accomplish. Psychotherapists are divided, it seems, into those who advocate doing almost nothing except asking questions (rational-emotive therapy), asking almost no questions (client-centred therapy), or asking some questions some of the time (eclectic therapy), but since questions relate to different underlying purposes it is not clear which, if any, strategy is the most effective. Those who write questionnaires and surveys for a living apparently spend their whole professional lives re-writing questions in an endless quest for one question which will not be either misunderstood, insulting or irrelevant. So far they haven't found one. The chapter on survey questions, however, is potentially the most useful in the whole book and I would recommend it to any potential author of a questionnaire.

The joke at the end of this longish shaggy-dog story is that if you really want people to tell you about their beliefs, thoughts and feelings you shouldn't try asking them questions at all. There is some research and it suggests that, from preschoolers to octogenarians, people will supply more information in response to statements, restatements, social "fillers" or even deliberate silences, than they will in response to questions. Is that clear?

RICHARD FOX

REFERENCE

POLLARD, A. and TANN, S. (1987) *Reflective Teaching in the Primary School*. London. Cassell.

GEEN, R. G. (1989). *Human Aggression.* Milton Keynes: Open University Press, £27.50 hbk., £9.99 pbk. ISBN 0-335-15346-1, 0-335-15345-3.

This is an intelligent little book which examines (with minimal use of jargon) the conceptually difficult phenomenon of human aggression. Dr Geen has written this book primarily for an undergraduate audience with a grounding in psychology and reasonably concentrates on *affective aggression* (that which is apparently accompanied by "anger") rather than *instrumental aggression* (that behaviour which serves some other and or purpose and does not have a strong affective component). The rationale is that much more is known about the former than the latter. The book provides a social psychological approach to human aggression but also sensibly contains contributions from physiological, developmental and personality psychology. Dr Geen's book is intended to provide an introduction to what is known about human aggression rather than recommend strategies for the control of elimination of aggression. In this sense, it succeeds remarkably well.

Human Aggression considers the varied approaches to the study of a clearly highly heterogeneous group of activities and advocates (reasonably) the involvement of "intent to harm" and "expectancy" in such phenomena. The author's approach is a deliberately simplified two-factor one in which the "state of the person" who is likely to aggress and the "situation that elicits behaviour" are considered to be of equal importance in the equation. In the case of the former tendency, the roles of genes, brain mechanism and body chemistry are briefly (but reasonably accurately) reviewed. Factors such as social learning (including modelling and imitative learning and reward) are also considered, as are the impact of social and cultural antecedents (subcultures and gangs).

The frustration — aggression hypothesis of Dollard *et al.* (1939) receives a substantial re-airing and some updating (including much evidence derived from teaching situations, "stress", negative affect, coping, norm violation etc.) in the section on interpersonal antecedents. This chapter is likely to be of great interest to those involved in education. A chapter on environmental antecedents deals with factors such as crowding, noxious odours and noise. Educators are also likely to be of the mass media (e.g., television and pornography) in certain forms of human aggression.

Human Aggression also has sections on intervening processes (such as depersonalisation, guilt and empathy), the role of individual differences (e.g., sex differences, cognitive development and personality variables) and hostility, health and adjustment (including material on catharsis and reinforcement).

I think that this is an excellent book which can also be read by educational professionals who could benefit from a relatively non-partisan view of the nature of human aggression and a sensible and fair evaluation of the conflicting claims concerning the genesis of such phenomena. Certainly, there are lots of myths which need to be exploded and Dr Geen goes some way towards countering some of the less supportable claims.

PAUL F. BRAIN

HARRIS, P. (1989). *Children and Emotion: the Development of Psychological Understanding.* Oxford: Basil Blackwell, £9.95 pbk. ISBN 0-631-16753-6.

EISENBERG, N., and MUSSEN, P. H. (1989). *The Roots of Prosocial Behaviour in Children.* Cambridge: Cambridge University Press, £8.95 pbk. ISBN 0-521-33771-2.

Children and Emotion examines the development of children's understanding of their own and others' emotional experiences. It is therefore part of an important movement within developmental psychology that is concerned with how children develop concepts of mental life or, more grandly put, a theory of mind. Paul Harris refers to Craik's (1943) paper as justification for the cognitive psychological perspective he adopts. Craik suggested that we test reality against models which represent it, and Harris argues that this is what children learn to do in relation to emotional experiences. In particular, he suggests that we use our imaginations to simulate the intentions or emotions of others and it is the increasing accuracy of this simulation with increasing age which explains many of the phenomena observed.

This thesis is pursued through a number of quite varied chapters, beginning with an account of the understanding of facial expression in infancy, progressing through chapters examining the literature on hurting and comforting, pride, shame and guilt, controlling emotion, and autism. The author picks his way through a complex literature, guided by his own theoretical model, with exemplary clarity. He has produced an elegant and scholarly book which deserves to be widely read by both academics and practitioners.

The Roots of Prosocial Behaviour in Children is an update of a book which the authors published in 1977 that reviewed the then limited amount of work on the topic. Much work has been carried out over the last 12 years which justifies the need for a new attempted synthesis. Prosocial behaviour is defined as referring to "voluntary actions that are intended to help or benefit another individual or group of individuals". The authors examine cultural or biological influences on such behaviour and there are chapters on cognitive and emotional factors associated with it. The book is necessarily broad-ranging, passing over social learning and Piagetian theory, role modelling, and moral behaviour. Indeed, in their introduction Eisenberg and Mussen quote from their earlier work to the effect that troublesome problems of methodology remain, studies contradict one another, and it is difficult to synthesise research findings and arrive at general conclusions.

The present book does end with a suggested model of prosocial behaviour: 19 boxes linked with arrows indicating direction of flow. This is, of course, no more than a set of categories and I have yet to come across a psychologist who finds other people's flow charts useful! This book is, then, a valuable source for the literature, but the uninitiated may finish more confused than enlightened.

Whilst *Children and Emotion* bears all the hallmarks of an author who has experienced an educational system which largely progresses its students according to their capacity to write a good essay, *The Roots of Prosocial Behaviour* comes from a tradition which is more information-based and categorical in its thinking. I must confess to a sneaking and perhaps ungrateful regard for a latter tradition, since with the former one is never quite sure whether one is carried by the argument or deceived by the prose style.

Much more seriously, however, is the alarming lack of contestants in psychology these days — with the decline of theories and the rise of cognitive psychology, theoretically-driven disputes have given way to the fine-tuning of models. Only the conceptual framework in which the models operate remains unexamined.

NEIL BOLTON

NEWMAN, D., GRIFFIN, P., and COLE, M. (1989). *The Construction Zone: Working for Cognitive Change in School.* Cambridge: Cambridge University Press, £8.95 pbk. ISBN 0-521-38942-9.

The core of this book is a set of analyses of children's developmental changes during classroom lessons and individual tutorials designed to teach basic concepts in such diverse areas as natural science, social studies, and arithmetic. The authors undertake four main tasks: to study students who are low performers in school, to illustrate how differences in performance often relate to differences on how children interpret school tasks; to examine the differences between educational and research practice and suggest how the interaction of the two can yield important insights about the process of cognitive change; to study methods for analysing cognitive change as it occurs in a variety of settings, and, finally, to construct a theory of cognitive change. These tasks recur in different contexts and with different levels of analyses.

The book's central contribution is to focus on the "instructional interaction as the locus of the constructive activity which constitutes cognitive change" (p. 4). The authors' research strategy evolved from earlier attempts to specify how cognitive tasks and behaviours vary across social contexts. They examine how best to adapt techniques from microsociology, experimental psychology and ethnographic studies of the classroom, to describe and explain cognitive change as a process of joint constructive interaction. They retain a Piagetian development perspective which treats intellectual change as an individually constructed process while making it clear that developmental change is simultaneously a social process of cultural transformation as emphasised by Vygotsky and his students.

The project on which the book is based was conducted in its first year with third- and fourth-grades in a team-taught classroom, and, in the second year with fourth-grade classrooms. Eighty children were sampled over a period of two years. Teachers were involved in designing seven curriculum modules, each of which was intended as a formative experiment in creating the same task across a variety of contexts. The topics investigated ranged from science, e.g., electricity, animals, household chemicals, to division in mathematics, to mapping and native American Indians in social studies, to a unit on memory and study skills. Striking differences emerged between the teacher and researcher perspectives which became a topic in itself.

The approach outlined in the book is a concrete example of the principles of cognitive change. The findings reported are as much revisions and re-interpretations of the theory based on events in the classroom as a presentation of empirical results: it is simultaneously theoretical and empirical.

The authors explain in a challenging way why activity — represented as tasks — and situations are integral to cognition and learning. The notion of a construction zone is the locus of social negotiations about meanings, and it is, "in the context of schools, a place where teachers and pupils may appropriate one another's understandings" (p. xxi). They deny that conceptual knowledge can be abstracted from the situations in which it is learned and used. They make deliberate use of the social and physical context to understand learning and cognition. In four of their chapters they take up specific cases of classroom activities on which to build the case that cognitive change is a process of active construction, observing how children's cognitive processes change as teacher and student and the students themselves interact. Their central argument, which is persuasively presented, is that for cognitive change inclusion of the social environment is a necessity.

A key concept for explaining how cognitive change can happen is based on a distinction made by Leontev which replaces Piaget's notion of assimilation with the concept of "appropriation". Children appropriate culturally devised "tools" through involvement in culturally organised activities in which the tool plays a role, for example, the mother tongue, systems and operations used on numbers in schools, businesses and homes. Children have only to come to an understanding that is adequate for using the culturally elaborated object in the novel circumstances they encounter. This process in which teachers appropriate something a child does, interpreting it in terms of their own frames of reference, is a means by which cognitive change can take place.

They examine some of the important implications of their theory for assessment, claiming that assessment while teaching differs in at least three ways from the usual practice of dynamic assessment: the adult's priorities are reversed so that instruction is not subordinated to assessment; unable to enjoy the luxury of working one-to-one with students, the teacher can, in small group settings, sample the apparent match between the children's behaviour and the behaviour she expects as a part of her conception of the task; and, remind us that many instructional units do not decompose themselves into a neat sequence of levels to be mastered in an invariant sequence with a single correct route to mastery. Appropriate child behaviours come in great variety, requiring flexible expertise on the teacher's part to weave them into a productive instructional interaction.

In stressing social interaction and collaborative learning the authors persuasively demonstrate how teaching through appropriation within a construction zone can accommodate a new view of knowledge and learning. It offers, too, a timely re-appraisal of the assumptions about learning underlying current classroom practice and teacher education programmes. This book will be influential in helping teachers and researchers to try to frame a convincing account of the relationship between explicit knowledge and implicit understanding.

Gareth Harvard

Ochse, R. (1990). *Before the Gates of Excellence: the Determinants of Creative Genius.* Cambridge: Cambridge University Press, £9.95 pbk. ISBN 0-521-37699-8.

The past decade has had some bumper years for insightful books on creativity. Before that we had Arthur Koestler's contribution, brilliant in its way but not to everyone's taste, some interesting but rather insubstantial volumes by authors such as Peter McKellar and Liam Hudson, and a number of bold but largely unsuccessful efforts to show that creativity is a measurable human trait that could be assessed by "creativity tests" which were expected to provide useful predictive instruments. (They didn't, and the articles on creativity measurement which at one time threatened to fill the educational journals have all but disappeared.) Since the late seventies we have seen David Perkins' highly successful attempt to debunk and demystify, and a shorter but effective book in the same vein by Robert Weisberg, as well as Howard Gruber's marvellous exploration of the mind of Charles Darwin, some insightful volumes by writers such as Mihaly Csikszentmihalyi, Howard Gardner, Robert Albert, David Feldman and Dean Simonton, and Teresa Amabile's experimental work demonstrating the negative influences of certain kinds of motivation on creative activities. Moreover, a number of workers in other fields, including Robert Sternberg and Herbert Simon, have also made useful contributions to our understanding of creativity. Writers such as the sociologist Augustine Brannigan have ensured that the field has not been left entirely to psychologists.

Few authors of books on creativity are noticeably modest. Most often, they are keen to tell a good story and advance their brand of wisdom, and persuade the reader to share their pet theory. It sometimes appears that the factual evidence is being carefully selected or even neglected, in the interest of helping the author's efforts to push the particular line that he or she wants readers to accept. In contrast, R. Ochse, the pseudonymous author of *Before the Gates of Excellence*, has written a book that is remarkably self-effacing. She rarely allows her own point of view to intrude, and is generally content to survey the work of others. It was something of a novelty for this reviewer to find himself wishing that more room had been allowed for the author's own views and opinions. If the idea of a students' textbook on creativity is not a contradiction in terms, Ochse comes very close to providing one.

The book is not without faults. For instance, there is too much dependence on the not always reliable "potted biographies" which the diligent Catherine Cox assembled almost three-quarters of a century ago. The somewhat pointless exercise in which Cox and other colleagues of Lewis Terman estimated the IQs of dead people is solemnly and uncritically reported. The author's account of cerebral localisation and its implications appears naive. Her discussion of the role of unconscious processes fails to take account of modern evidence concerning the very limited role of conscious thought in many aspects of cognition. And she quotes, with apparent approval, a number of statements that are nonsensical, such as "The conscious mind reaches the state of pure consciousness, which is the source of all thinking."

All the same, as a survey of recent contributions to our understanding of creativity, *Before the Gates of Excellence* is generally well-written and well-informed, as well as being reasonably comprehensive. It is a very useful book to have. And since the paperback version is commendably inexpensive — well done Cambridge University Press: if only certain other publishers would follow your example! — Ochse's book should become a valuable resource for students and others who would like to have a better understanding of the origins of creativity.

Michael J. A. Howe

Radford, John, (1990). *Child Prodigies and Exceptional Early Achievers*. Hemel Hempstead: Harvester Wheatsheaf, £30.00 hbk., £9.95 pbk. HB ISBN 7108-11403.

This excellent book surveys, discusses and evaluates a wide range of evidence about child prodigies and exceptional early achievers. It is maturely informed, scholarly, up to date, and yet written in attractively non-technical language. It has no axe to grind other than to display the development of human individuals and how they come to show such a variety of achievements and enthusiasms. It bulges with human interest and opens our eyes to the wealth of influences that shape the courses taken by young people's lives. In short, the book is both entertaining and instructive.

The lively tour on which Radford conducts us is divided into 11 chapters with titles such as: the marvellous child, top of the class, sports and games, calculators and mathematicians, precocious artists, and young leaders. There is an extensive bibliography and a full index. I regret that a book of this inviting kind could not have been available to me 40 years ago when I was in the early stages of a career in psychology.

In reviewing this book I shall not attempt to summarise its eclectic contents but shall try to convey the general flavour by quoting parts of the final chapter. This chapter, entitled "The Problem of the Prodigy", touches some issues raised earlier in the book. "There are," says Radford, "three main groups of questions. One is concerned with explaining prodigies, whatever one likes to mean by "explain". This would include the circumstances under which they appear, the physiological, psychological and social factors that seem to be present and the experimental or mental processes that characterise them. A second is concerned with what can be learned from them about wider groups, or about psychological processes in general. And a third is concerned with matters of prognosis, i.e. what happens to prodigies later on; and with the corollary, whether great talent manifests itself early" (p. 218).

There are also three other sorts of questions. "One is questions of technology; how to bring about a given result, whether or not we can explain the process — in this case, for example, produce prodigies. Another concerns policy, what ought to be done, how the very highly gifted should be treated. The last sort is philosophical and metaphysical, concerned with what sort of being the human is, and how it may be investigated. All of these have been touched on, not so much deliberately as because they arise inevitably" (p. 218).

Radford then mentions his own long-term view that human behaviour is a proper, indeed essential subject for the methods of science. "This involves considering, though not uncritically, data from any source. Ideally, it involves experiment, failing that observation and compilation, in any case reason and reflection. These show, it is held, something of the extraordinary variety, and the still greater potential, of the human species. They show too some of the ways in which potential can be thwarted or perverted, though in this particular book we have been more concerned with achievement than failure. Moreover, something at least is understood of how these processes work and how they might be made better — very far from all, but still greatly more than even a generation ago" (p. 219).

Throughout the book, Radford is explicit that he does not aim to construct any neat theories or to test any abstract hypotheses, and this may disappoint some readers. What he aims to do is to gather together, and make available for consideration, a wealth of real-life material drawn from various sources. This aim is achieved and has value enough to commend the book to a wide readership. For example, the material exposes as false many of the ideas that are popular about "talent" and "ability", and alerts theorists to facts that their theories may be missing. For applied psychologists who deal with individual children, the book provides a rich background of case material for purposes of comparison and guidance. The book is also relevant to cognitive scientists who study the acquisition of expert knowledge and skill. Finally, although the book is mainly about young people, it raises issues relevant to development throughout the lifespan and, indeed, relevant to the interactions between psychology and the literary enterprise of writing biographical stories about people's lives. In conclusion, this book is readable and enjoyable without recourse to journalistic licence. It clears much empirical ground and widens horizons. I recommend it highly.

IAN M. L. HUNTER

SPENCER, C., BLADES, M., and MORSLEY, K. (1989). *The Child in the Physical Environment: the Development of Spatial Knowledge and Cognition.* Chichester: John Wiley and Sons, £32.95 hbk. ISBN 0-471-91235-2.

Although spatial abilities tend to be under-represented in mainstream cognitive psychology, a large and flourishing field of research has developed in past years under the heading of *environmental cognition*. The literature in this field is contributed to by workers in a number of disciplines, including geography and urban planning as well as psychology. This book reviews the literature on the development of environmental cognition, no mean task, since the literature here is probably even more diverse than that relating to adults.

The book is divided into three sections. The first sets up the main issues in the area by reviewing theory and data on adult environmental cognition. The second and longest section looks at spatial development from infancy through to adolescence, and the third section deals with applied research.

The opening chapter indicates the breadth of the area and the range of investigative techniques employed within it. Environmental cognition is not just about how people find their way around (although that is a very important part of it), but includes topics less amenable to controlled experimental investigation, such as emotional judgments about places and the effect of different physical environments on people's well-being.

Chapter 2 reviews literature and relevant theory regarding adults' movement and orientation in space and also includes sections on human environmental needs and on images of distant places. Data on relating space to self and mental rotation are reviewed and also relevant theories such as J. J. Gibson's. The authors, however, are critical of Gibson's theory, pointing out that continuous visual monitoring is not needed for continued orientation in space during movement, a finding which they claim Gibson's theory cannot accommodate. I think this may not be entirely fair to Gibson, since his theory was about perception, and the specific claim was that perceptual information was sufficient to support movement through the environment and awareness of position in it. It is thus not incumbent on the theory to explain cases of environmental cognition when perceptual support of this sort is absent or interrupted. Gibson's theory does not rule out cognition but rather claims that for many everyday activities cognition is unnecessary.

These chapters set the scene for the main section of seven chapters on development of environmental cognition. Chapters 3 and 4 deal with development of spatial reference systems and coding of positions relative to self and external space. The authors pay particular attention to the literature on the relation between active movement and spatial orientation, one of the central foci of current research in this area, and this leads on to a good section on the development of way-finding and position updating during movement. The literature here is well reviewed, but the reader gains the impression that the authors are not too favourably impressed with developmental psychology as a discipline. For instance, they suggest that the non-inclusion of five infant subjects due to fussiness constitutes a "cavalier attitude towards... data". But those who work with infants know that this sort of subject loss is inevitable, and there seems nothing cavalier about eliminating subjects from a study when they are clearly likely to have been inattentive or otherwise not performing well.

The remaining five chapters in this section cover different aspects of the child's understanding of geographical space. Some of the authors' own work is described here in which young children go on walks with an adult through a novel urban environment, first with adult as guide and later reversing roles with them as guide. This nice technique leads to rich data on children's route learning and landmark use, derived both from the child's accuracy and their verbal reports.

This leads on to consideration of children's cognitive maps. Again, the authors focus on methodology, starting with a contrast between laboratory-based and "real life" approaches. They rightly suggest that laboratory-based investigation is not enough in itself, although I think they could have done more to show how laboratory studies are often needed to supplement ones done in real environments, where control over confounding variables is often non-existent.

A short chapter on how spatial representations develop with experience of the space leads into a chapter on children's use of maps. Again, this section contains some of the authors' own work on map use by young children, work which shows that preschool children have some ability to use a simple map to navigate a relatively simple route. In addition, they show that four-year-olds are capable of using grid references if these are made "child friendly".

The section is completed with a move to even larger scale spaces, looking at children's images of distant places. This involves a shift towards different methodology, since much of the work here focuses on children's judgments of the likability of different places and other work looks at the role of the media in influencing children's judgments of other countries.

The section on applications of research starts with a substantial chapter on the child with special needs. Much of this chapter focuses on visually handicapped children. There is a rich literature here that probably has as many implications for theory as it has for application, since the impressive spatial abilities of the blind confront us with the knowledge that at very least there must be alternative ways of orienting in space that are not dependent on visual input.

The final two chapters are on planning for children's needs and on spatial education. These chapters indicate the multi-disciplinary nature of the topic as well as its breadth, and the authors integrate these subtopics well with the rest of the book.

To sum up, this is an invaluable source book for anyone wishing to gain a good knowledge of the literature on environmental cognition. Probably inevitably, it is fairly hard reading and there is often a slant towards methodology that tends to distract from the development of a rich account of the growing child in the environment. In addition, I did not come away with a feeling that the developmental processes involved in children's changing environmental cognition had been explicitly addressed. This, however, is asking rather a lot, since one of the main failings of developmental psychology is the lack of research that directly deals with the processes of developmental change. So that is something for the future and, given the current state of the literature, this book is to be recommended as a very thorough treatment of a broad and often tricky area.

GAVIN BREMNER

STORFER, M. D. (1990). *Intelligence and Giftedness: the Contributions of Heredity and Early Environment*. San Francisco: Jossey-Bass, £35.00. ISBN 1-5542-185-7.

Miles Storfer has set himself the enormously difficult challenge of attempting to survey the vast and diverse body of evidence that exists concerning the causes of differences between children in intellectual ability. He is not entirely successful, but in view of the magnitude of the task that is by no means a damning criticism. If Storfer has not been able to weld all the research findings he has examined into an account which is particularly cohesive or consistently critical and well-informed, it is doubtful whether any other single author could have done much better.

This is a long book, with the bibliography alone requiring 70 pages. There are 20 chapters in total, and the first six of them are essays which consider broadly demographic forms of evidence concerning a number of factors that may influence childhood IQ. Thus one chapter surveys studies of identical twins, another examines the effects of adoption, another considers IQ resemblances within families, and there are chapters on race-related differences and on evidence concerning the causes of rises in average IQs. With such a broad canvas there are inevitably some omissions and some failures to take a sufficiently critical approach to the data. For instance, Chapter Two, which examines the effects of birth order, parental age, and number of siblings on IQ, gives the important work of Marjoribanks and Zajonc only the briefest of mentions. And when Storfer discusses a number of interesting findings which show that child IQ is correlated (positively) with variables such as maternal age and home discipline, he pays too little attention to the possible hazards involved when one draws causal inferences from correlational data.

Chapters Seven to Nine contain a substantial amount of recent evidence about early precursors and predictors of later ability differences. Infant habituation studies have provided some of the most fascinating findings. For instance it is clear that measures of habituation in the first six months are reasonably good predictors of childhood IQ, and it also appears very probable that early habituation is itself strongly influenced by the mother's actions in the earliest months. Consequently, there now exists a fairly strong case for claiming that the ways in which the mother interacts with her child in the first few months of life may have important consequences for later intellectual development. Again however, the evidence is largely correlational, and Storfer fails to give sufficient weight to the problems that beset any author who wishes to make inferences concerning cause-and-effect relationships. Yet it is clear that even very young infants are sensitive to various kinds of maternal stimulation, and well before the child reaches six months of age effective mothers can elicit chains of mutually responsive mother-infant interactions which are very likely to lead to the infant's gaining skills that can facilitate future learning.

Later chapters provide rather sketchy surveys of research studies that have examined the possibility of stimulating early cognitive development. There are brief descriptions of some exciting research by William Fowler and by Grover Whitehurst, both of whom have demonstrated that it is indeed possible to accelerate rather considerably the language development of "normal" children. Towards the end of the book Storfer introduces a number of (to this reviewer) unconvincing views concerning biological determinants of ability differences. He advances accounts of the workings of hemispheric localisation which I believe are considerably simpler and more clear-cut than are justified by any hard evidence. Equally unconvincing, in my view, are his efforts to persuade the reader to accept a kind of neo-Lamarckian theory whereby the potential achievements of a new generation may be increased by what the previous generation has learned. I suspect Storfer has not exposed himself to Vygotsky's accounts of the sheer potency of cultural influences, or to Charles Super's evidence that large differences in motor development between babies of different races, which have long been assumed to be intractable and to provide firm evidence of underlying biological causes of variability, can be wiped out by cultural change in a single generation. Another shortcoming is that whilst insisting on providing a quantitative figure for of the importance of hereditary causes of intelligence (a step which many geneticists would consider invalid) Storfer has very little to say about possible mechanisms whereby genetic influences might actually make their presence felt. Since the practical implications of any demonstration of genetic causation depend upon the manner in which such causation operates, which largely determines the extent to which matters can (if it is desired) be altered or remedied by appropriate environmental modifications, this omission is regrettable.

Undoubtedly, this book has serious faults. Even as a survey of the evidence it is sometimes too selective, omitting important findings, and it is often insufficiently critical of the research and the ideas that are reported. But faults and failures are inevitable when someone attempts the impossible, as Storfer has done here. He has had the courage to attempt a task that others have found too daunting. Miles

Storfer's valiant effort to bring order to a mass of detailed evidence has at the very least given us an invaluable survey of research into some of the many influences that contribute to individual differences in intellectual development.

MICHAEL J. A. HOWE

WRAY, D., BLOOM, W., and HALL, N. (1989). *Literacy in Action. The Development of Literacy in the Primary Years.* Basingstoke: Falmer Press, £8.95 pbk. ISBN 1-85000-605-9.

The current popularity of action research in promoting reflection on classroom procedures has encouraged educational writers to be more imaginative in the way theory and practice are related and presented. *Literacy in Action* is a refreshing addition to an ever increasing range of material on literacy teaching. Written as an introduction to literacy education from nursery school to the end of junior school, the book raises important issues in a simple, yet effective way and is likely to stimulate lively interest and debate amongst students new to the field. The three authors, David Wray, Wendy Bloom and Nigel Hall, are already well known for their contributions to the literature on literacy teaching and for their knowledge of children and classrooms. Together they have produced what they claim to be, an interactive book which aims "to explore ways in which teachers go about developing the literacy of children in their charge". Their whole philosophy may be summed up in the following words: "The adults who are most useful to children are those who *facilitate* the development of literacy rather than *instruct* children in it."

The book follows up this idea by facilitating understanding about literacy teaching. It uses activities which encourage the reader to participate and respond. Each chapter includes a number of observational strategies for collecting information and a series of points to help the reader make sense of their data. The observations are appropriate and varied, providing a range of simple strategies which may be used in a variety of classroom contexts. The "points for consideration" help the reader to reflect and to appreciate the discussion which follows.

Although the authors do not intend to provide a comprehensive consideration of the theories behind literacy learning, the material is sited within a theoretical framework which reflects the totality of the culture in which literacy experiences are embedded. The work of Heath (1983) and de Castel, Luke and Egan (1986) provides evidence of the socio-cultural influences on young children's ability to make sense of literacy within the social settings in which it is experienced. These influences are explored in the first chapter which shows the range of print possibilities within a literate society and raises the question of how far the school environment may be made more reflective of the literacy encounters children may meet outside the school. The second chapter goes on to explore a classroom scenario and analyses the messages of print within it. These two chapters set out a contemporary picture of literacy experiences both in society and in school but it was a pity that this did not include some consideration of the literacy encounters of different ethnic groups. I looked in vain for an example which might have involved bilingual children or focused on appropriate resource materials. In an introductory book which is establishing attitudes and approaches, such an omission seems a lost opportunity. The rest of the book is organised to enable the student reader to appreciate the literacy opportunities at different ages and includes sections on the beginnings of literacy, literacy in the infant school years, the early junior school years and the later junior years.

One of the excellent features of this book is the able way in which current theoretical insights are incorporated in the discussion alongside concrete examples. In the later chapters this is not so apparent, which probably reflects recent research which has tended to focus on the literacy in the early years. What the chapters on the junior years do very well is to emphasise the range of literacy opportunities which might occur across the curriculum. Practical observations and strategies for developing literacy competence are clearly described and evaluated.

For the reader who is unfamiliar with the field this is a valuable starting point. The annotated bibliography cites a number of useful texts, but one cannot help wishing that the authors had included further reading suggestions for those who wished to follow up some of the issues raised. Perspectives on language and literacy as reflected in the work of Goelman (1984), Street (1984), Cook-Gumperz, J. (1986), Meek and Mills (1988), and Garten and Pratt (1989) would have acknowledged the wealth of theoretical discussions which have informed much of their competent introduction to the field.

The book is accessible and enabling. For the audience for which it is intended, it will have much to offer.

Gill Beardsley

THE BURT AFFAIR

Letter to the Editor

Sir,

One would not usually wish to reply to a book review, but the Professors Clarke, in their remarks about *The Burt Affair* (Joynson, 1989), fall so far short of what may reasonably be expected that some response is required.

A review ought to offer the reader an indication of contents, and a balanced evaluation. It is only towards the end of their remarks that the Clarkes (Clarke and Clarke, 1990) find space to mention the book at all. Even then they make no attempt to outline its contents, merely dismissing it out of hand as containing "many errors of fact". The sole evidence they offer is as follows: "he paraphrases our statement 'Burt probably had little if any systematically collected MZA data'" (Clarke and Clarke, 1990, p. 123). But even in this solitary instance, it is the Clarkes who are in error. My statement was not a paraphrase of their statement. It was a summary, in my own words, of what I took to be their view as expressed in a number of remarks in the two sources mentioned. From the former I quoted their statement that there is "no evidence that Burt gathered data for himself" (p. 18), which refers to twins in general and not MZA twins only. They ignore this. From the latter I might have quoted: "... it seems possible, even probable, that Burt himself never had any systematic data of his own at all" (p. 173); and I might have added that, referring to Burt's list of items left at UCL, they say, "The most important and unique, on twins, is never mentioned, nor does it seem to have been transported to Aberystwyth. Why not? Almost certainly because it did not exist" (ibid). Elsewhere referring to the same list, they write that there is "no mention of twin data nor any other material relating to intelligence test results for relatives" (Clarke and Clarke, 1979: 151) — with the obvious intention of implying that Burt had no such data.

It is also true that the Clarkes were especially concerned about MZA data; and I noted this point (1989, 168-169). But it will be seen that the Clarkes did not limit their doubts to MZA material, nor even to MZ material. The passage they choose to extract was only part of a general scepticism about Burt's kinship data. Taken in isolation, it gives a misleading idea of the views, they expressed then, however much they may regret them now.

Readers of my book (especially pp. 243-248, 284-294, 302-303 and 316-320) will quickly discover why the Clarkes say so little about its contents, and why their evaluation is so unbalanced.

Sincerely,

R. B. Joyson

REFERENCES

Clarke, A. D. B., and Clarke, A. M. (1979). The cardinal sin. Review of Hearnshaw (1979). *Nature* 282, 150-151.

Clarke, A. D. B. and Clarke, A. M. (1980a). Comments on Professor Hearnshaw's *Balance Sheet on Burt.* In Beloff, H. (Ed.), A Balance Sheet on Burt. Supplement to *Bulletin of the British Psychological Society* 1980, 33, 17-19.

Clarke, A. D. B., and Clarke, A. M. (1980b). Review of Hearnshaw (1979). *Br. J. Psychol.*, 71, 172-173.

Clarke, A. D. B., and Clarke, A. M. (1990). Review of Joynson (1989). *Br. J. educ. Psychol.*, 60, 122-124.

Hearnshaw, L. S. (1979). *Cyril Burt: Psychologist*, London: Hodder and Stoughton.

Joynson, R. B. (1989). *The Burt Affair*, London and New York: Routledge.

PUBLICATIONS RECEIVED

Books for review have been selected from the following
("n.p." means "no price mentioned")

Barling, J. (1990). *Employment, Stress and Family Functioning*. Chichester: John Wiley, pp. 286, £39.00 hbk. ISBN 0-471-91773-7.

Black, H., Devine, M., and Turner, E. (1989). *Aspects of Assessment*. Edinburgh: SCRE, pp. 69, n.p. pbk. ISBN 0-951-2792-62.

Bruner, J., and Haste, H. (1987). *Making Sense*. London: Routledge, pp. 204, n.p. pbk. ISBN 0-416-92490-5.

Bryman, A. and Cramer, D. (1990). *Quantitative Data Analysis for Social Scientists*. London: Routledge, pp. 290, £10.99 hbk. ISBN 0-415-02664-4.

Butterfield, G., and Bryant, P. (1990). *Causes of Development*. Hemel Hempstead: Harvester Wheatsheaf, pp. 266, £35.50 hbk. ISBN 0-7108-1140-3.

Chessum, L. (1989). *The Part-time Nobody*. Bradford: Wycrow/University of Bradford, pp. 90, n.p., pbk.

Chiland, C., and Young, J. G. (1990). *L'Enfant dans sa Famille: Le Refus de L'Ecole un Apercu Transculturel*. Paris: Presses Universitaires de France, pp. 268, 198FF, pbk. ISBN 222660796-90.

Dijkstra, S., Van Hout Wolters, B., and Van Der Sijde, P. (1989). *Research on Instruction*. Englewood Cliffs, N.J.: Educational Technology Publications, pp. 184, $39.95 hbk. ISBN 0-87778-221-0.

Docking, J. (1990). *Education and Alienation in the Junior School*. Basingstoke: Falmer Press, pp. 235, £10.95 pbk. ISBN 1-85000-572-9.

Dornbusch, S. M., and Strober, M. H. (1990). *Feminism: Children and the New Families*. Hemel Hempstead: International Book Distributors, pp. 366, £14.50 pbk. ISBN 0-89862-514-9.

Duveen, G., and Lloyd, B. (1990). *Social Representations and the Development of Knowledge*. Cambridge University Press, pp. 191, £25.00 hbk. ISBN 0-524-36368-3.

Firth, H. and Rapley, M. (1990) *From Acquaintance to Friendship*. Kidderminster: BIMH Publications, pp. 222 £14.75 pbk. ISBN 0-906054-71-0.

Foot, H., Morgan, M., and Shute, R. (1990). *Children Helping Children*. Chichester: John Wiley and Sons, pp. 392, £49.95 hbk. ISBN 0-471-92292-7.

Fraisse, P., and Segui, J. (1990). *L'Année Psychologique*. Paris: Presses Universitaires de France, pp. 160, n.p. pbk. ISBN 2-13-043-085-6.

Fraser, W. I. (1990). *Issues in Mental Retardation Research*. London: Routledge, pp. 506, £45.00 hbk. ISBN 0-415-013630-1.

Frye, N. (1990). *Anatomy of Criticism*. London: Penguin, pp. 383, £7.99 pbk. ISBN 0-14-012480-2.

Furth, H. (1990). *Knowledge as Desire*. New York: Columbia University Press, pp. 179. $16.00 ISBN 0-231-06459-4.

Goldstein, S., and Goldstein, M. (1990). *Managing Attention Disorders in Children*. Chichester: John Wiley and Sons, pp. 451, £39.30 hbk. ISBN 0-471-61137-9.

Hall, E., Hall, C., and Leech, A. (1990). *Scripted Fantasy in the Classroom*. London: Routledge, pp. 170, £30.00 hbk. £9.99 pbk. ISBN 0-415-04811-7 and 0-415-04812-5.

Hinde, R., Perret-Clermont, A. N., and Stevenson-Hinde, J. (1987). *Social Relationships and Cognitive Development*. Oxford: Oxford University Press, pp. 384, £25.00 pbk. ISBN 0-19852167-7.

Hoffman, L., Paris, S., Hall, E. and Schell, R. (1988). *Developmental Psychology Today: Fifth Edition* London: McGraw-Hill, pp. 616, n.p. hbk. ISBN 0-07-55027-X.

Hurlock, E. (1989). *Child Development: Sixth Edition*, London: McGraw-Hill, pp. 570. n.p. pbk. ISBN 0-07-Y8531-6.

Jones, N., and Frederickson, N. (1990). *Refocusing Educational Psychology*. Basingstoke: Falmer Press, pp. 224, £10.95 pbk. ISBN 1-85000-492-7.

Johnson-Martin, N., Attermeier, S. and Hacker, B. (1990). *The Carolina Curriculum for Preschoolers with Special Needs*. Baltimore: Paul H. Brookes, pp. 334, £25.00 pbk. ISBN 1-55766-032-8.

Leone, P. E. (1990). *Understanding Troubled and Troubling Youth*. London: Sage Publications, pp. 317, £14.95 pbk. ISBN 0-8039-3281-2.

MacWhinney, B., and Bates, E. (1990). *The Cross-linguistic Study of Sentence Processing.* Cambridge: Cambridge University Press, pp. 489, £25.00 hbk. ISBN 0-521-261961.

Merrett, F., and Wheldall, K. (1990). *Positive Teaching in the Primary School.* London: Paul Chapman, pp. 111, £6.95 pbk. ISBN 1-85396-087.

Montemayor, R., Adams, G. and Gullotta, T. (1990). *From Childhood to Adolescence* London: Sage, pp. 308, £14.95 pbk. ISBN 0-8039-3404-1.

Neff, L. (1988) *One of a Kind.* Portland, Oregon: Multnomah, pp. 197, $7.95 pbk. ISBN 0-88070-232-X.

Newbould, S. (1990). *Mental Handicap Research.* Kidderminster: BIMH Publications, pp. 108, n.p. pbk.

Ochse, R. (1990). *Before the Gates of Excellence.* Cambridge: Cambridge University Press, pp. 300. £9.95 pbk. ISBN 0-521-37699-8.

Pardeck, J. (1989). *Child Abuse and Neglect: Theory, Research and Practice.* New York: Gordon and Breach, pp. 194. $39.00 hbk. ISBN 0-677-258860-7.

Presland, J. (1989). *Overcoming Difficult Behaviour.* Kidderminster: BIMH Publications, pp. 342, £19.95 pbk. ISBN 0-906054-67-2.

Quicke, J., Beasley, K., and Morrison, c. (1990). *Challenging Prejudice Through Education.* Basingstoke: Falmer Press. pp. 195, £9.95 pbk. ISBN 1-85000-693-8.

Radford, J. (1990). *Child Prodigies and Exceptional Early Achievers.* Hemel Hempstead: Harvester Wheatsheaf, pp. 255, £9.95 pbk. ISBN 0-7450-0751-1.

Schiff, J. L., Kadane, J. B., and Froman, C. (1990). *A Programme for Schizophrenia.* Upton-upon-Severn, Worcs: The Self-Publishing Association Ltd. pp. 87, £9.95 hbk. ISBN 1-85421-073-4.

Schewebel, M., Maher, C. A., and Fagley, N. S. (1990). *Promoting Cognitive Growth Over the Life Span.* Hillsdale, N.J.: Lawrence Erlbaum Associates, pp. 244, $39.95 hbk. ISBN 0-8058-0663-6.

Sigmon, S. B. (1990). *Critical Voices on Special Education.* New York: State University of New York Press, pp. 265, n.p. hbk. ISBN 0-78914-0319-X.

Sitterly, C., and Duke, B. (1988). *A Woman's Place.* Hemel Hempstead: Prentice Hall, pp. 365, £23.10 pbk. ISBN 0-13-961954-2.

Steedman, C. (1990). *Childhood, Culture and Class in Britain.* London: Virago, pp. 343, £16.99 pbk. ISBN 1-85381-123-8.

Stein, R. H. (1990). *Ethical Issues in Counseling.* New York: Prometheus Books, pp. 174, £22.95 hbk. ISBN 0-87975-557-1.

Storfer, M. D. (1990). *Intelligence and Giftedness.* San Fransisco: Jossey-Bass pp. 636, $35.00 hbk. ISBN 1-55542-185-7.

Sugarman, S. (1990). *Piaget's Construction of the Child's Reality.* Cambridge: Cambridge University Press, pp. 258, £10.95 pbk. ISBN 0-521-37967-9.

Tennant, M. (1990). *Psychology and Adult Learning.* London: Routledge, pp. 181, £10.99 pbk. ISBN 0-415-05032-4.

Terwogt-Kouwenhoven, K. (1990). *Niet Gewogen Toch Te Licht Bevonden.* Kampen: Mondiss, pp. 146, n.p. pbk. ISBN 90-5337-0021.

Tesch, R. (1990). *Qualitative Research.* Basingstoke: Falmer Press, pp. 330, £12.95 pbk. ISBN 1-85000-609-1.

Tufte, E. R. (1990). *Envisioning Information.* Cheshire, Connecticut: Graphics Press, pp. 126, $48.00 hbk.

Walkerdine, V. (1990). *The Mystery of Reason.* London: Routledge, pp. 230, £9.99 pbk. ISBN 0-415-05233-5.

Zebrowitz, L. A. (1990). *Social Perception.* Milton Keynes: Open University Press, pp. 231. £9.99 pbk. ISBN 0-335-09860-6.

Zimmerman, B. J., and Schunk, D. H. (1989). *Self-Regulated Learning and Academic Achievement.* Berlin: Springer-Verlag, pp. 212. DM 98,000 hbk. ISBN 0-387-9634-9.

AUTHOR INDEX TO VOLUME 60, 1990

ANDERSON, M.: see Johnston, R. S.

ANDERSON, R.: see Chadwick, D.

BANKS, J.: The Written Recall of Printed Stories by Severely Deaf Children 192

BARRY, J. F.: see Seddon, G. M.

BAUDOIN, C.: see Wallon, P.

BEHRENS, J. T.: see Kulhavy, R. W.

BEVERIDGE, M.: see Reid, D. J.

BORG, M. G.: Teachers' Perception of Primary Schoolchildren's Undesirable Behaviours: the Effects of Teaching Experience, Pupil's Age, Sex and Ability Stream 220

BUTLER, R.: When Pupils Do Not Understand the Determinants of their Success and Failure in School: Relations between Internal, Teacher and Unknown Perceptions of Control and School Achievement . 63

CAIRNS, E.: The Relationship between Adolescent Perceived Self-Competence and Attendance at Single-Sex Secondary School 207

CHADWICK, D.: The Examination Attainments of Secondary Pupils who Abuse Solvents 180

CHAPMAN, J. W.: Some Antecedents of Academic Self-Concept: a Longitudinal Study 142

CREESER, R.: See Plewis, I.

CUNNINGHAM, C.: see Sloper, P.

DADDS, M. R.: see King, N. J.

DE METSENAERE, M.: see Van Overalle, F.

DUFFY, L.: see Wishart, J. G.

DUNCAN, J.: see Norwich, B.

ESHEL, Y.: Individualised Instruction, Congruent Authority Structure and Student Outcomes 1

FALZON, J. M.: see Borg, M. G.

FOGELMAN, K.: see Hibbett, A. (first and second entries)

FYFE, R.: see Banks, J.

GOW, L.: see Kember, D.

GRAY, C.: see Banks, J.

GULLONE, E.: see King, N. J.

HIBBETT, A.: Future Lives of Truants: Family Formation and Health-Related Behaviour 171

HIBBETT, A.: Occupational Outcomes of Truancy 23

HOLLIGAN, C.: see Johnston, R. S.

HOUGHTON, S.: The Effects of Limited Private Reprimands and Increased Private Praise on Classroom Behaviour in Four British Secondary School Classes 255

INGRAM, J.: Varieties of Curricular Experience: Backmarkers and Frontrunners in the Primary Classroom . 52

JOHNSTON, R. S.: Perceptual Dysfunction in Poor Readers: Evidence for Visual and Auditory Segmentation Problems in a Sub-group of Poor Readers 212

JUKES, R.: see Houghton, S.

KEMBER, D.: Cultural Specificity of Approaches to Study 356

KING, N. J.: Student Perceptions of Permissiveness and Teacher-instigated Disciplinary Strategies 322

KNUSSEN, C.: see Sloper, P.

KULHAVY, R. W.: Response Feedback, Certitude and Learning from Text 161

KURMAN, J.: see Eshel, Y.

LAMBOURNE, R.: see Chapman, J. W.

LEACH, D. J.: Parental Involvement in the Teaching of Reading: a Comparison of Hearing Reading, Paired Reading, Pause, Prompt, Praise and Direct Instruction Methods 349

MANOR, O.: see Hibbett, A. (second entry).

MASON, J.: The Development of Concepts about Written Language in the First Three Years of School 266

MOLLOY, G. N.: see Pierce, C. M. B.

MOONEY, A.: see Plewis, I.

MOORE, D. W.: see Townsend, M. A. R.

NORWICH, B.: Attitudes, Subjective Norm, Perceived Preventive Factors, Intentions and Learning Science: Testing a Modified Theory of Reasoned Action 312

ORION, R.: see Butler, R.

PERRETT, D. I.: see Johnston, R. S.

PIERCE, C. M. B.: Psychological and Biographical Differences between Secondary School Teachers Experiencing High and Low Levels of Burnout 37

PLEWIS, I.: Time on Educational Activities at Home and Educational Progress in Infant School 330

REID, D. J.: Reading Illustrated Texts: a Microcomputer-based Investigation of Children's Strategies 76

RIDING, R. J.: see Thompson, S. V.

SÄLJÖ, R.: Problem-solving, Academic Performance and Situated Reasoning. A Study of Joint Cognitive Activity in the Formal Setting 245

SEDDON, G. M.: The Relative Effectiveness of Four Methods of Teaching Practical Scientific Skills to Secondary School Students 99

SHARPE, A.: see Houghton, S.

SIDDALL, S. W.: see Leach, D. J.

SILVA, P. A.: see Chapman, J. W.

SKAALVIK, E. M.: Experiences in School as Determinants of Behaviour in a Discussion Situation 133

SLETTA, O.: see Skaalvik, E. M.

SLOPER, P.: Factors Related to the Academic Attainments of Children with Down's Syndrome 284

STOCK, W. A.: see Kulhavy, R. W.

THOMPSON, S. V.: The Effect of Animated Diagrams on the Understanding of a Mathematical Demonstration in 11- to 14-Year-Old Pupils 93

THORNTON, N. E.: see Kulhavy, R. W.

TOWNSEND, M. A. R.: Headings Within Multiple-Choice Tests as Facilitators of Test Performance 153

TUCK, B. F.: see Townsend, M. A. R.

TURNER, S.: see Sloper, P.

VAN OVERALLE, F.: The Effects of Attribution-based Intervention and Study Strategy Training on Academic Achievement in College Freshmen 299

WALLON, P.: A New Criterium in the Assessment of the Behaviour of the Drawing Child 338

WHELDALL, K.: see Houghton, S.

WILTON, K. M.: see Townsend, M. A. R.

WINSTON, K. S.: see Kulhavy, R. W.

WINTER, S.: Teacher Approval and Disapproval in Hongkong Secondary School Classrooms 88

WISHART, J. G.: Instability of Performance on Cognitive Tests in Infants and Young Children with Down's Syndrome 10

WORRALL, N.: see Ingram, J.

WYNDHAMN, J.: see Säljö, R.

YULE, W.: see Chadwick, D.

CENTRAL LIBRARY

EDITORIAL ACKNOWLEDGEMENT

The Editor and Assistant Editors of *The British Journal of Educational Psychology* gratefully acknowledge the assistance of the following, who have refereed papers submitted between 1st April 1989 and 31st March 1990.

Dr R. Beard
Prof. S. N. Bennett
Dr B. Brown
Prof. G. Brown
Dr R. J. Cameron
Dr T. Cline
Dr A. Dobbins
Dr T. Dolan
Prof. J. Eggleston
Prof. N. J. Entwistle
Dr R. Fox
Dr J. Grey
Dr P. Gurney
Dr C. Harrison
Mr G. Harvard
Dr M. Howe
Dr M. Hughes
Dr C. Kyriacou
Ms F. J. Macleod
Prof. D. Macnamara
Dr J. Mills
Mr C. Mills
Dr P. Murphy
Dr P. Preece
Dr C. Rogers
Dr D. Shorrocks
Dr L. Smith
Mr B. Stratford
Dr K. Topping
Prof. G. Underwood
Prof. D. Warburton
Dr A. Webster
Prof. E. C. Wragg
Dr M. Youngman

Reading and Writing:

An Interdisciplinary Journal

Editor:

R. Malatesha Joshi, *Dept. of Reading, School of Education and Human Development, Fayetteville State University, NC, USA*

Reading and writing skills are fundamental to literacy. Consequently, the processes involved in reading and writing and the failure to acquire these skills, as well as the loss of once well-developed reading and writing abilities have been the targets of intense research activity involving professionals from a variety of disciplines, such as neuropsychology, cognitive psychology, psycholinguistics and education. The findings that have emanated from this research are most often written up in a lingua that is specific to the particular discipline involved, and are published in specialized journals. This generally leaves the expert in one area almost totally unaware of what may be taking place in any area other than his own.

Reading and Writing: An Interdisciplinary Journal cuts through this fog of jargon, breaking down the artificial boundaries between disciplines. The journal focuses on the interaction among various fields, such as linguistics, information processing, neuropsychology, cognitive psychology, speech and hearing science and education.

Reading and Writing publishes research articles, critical reviews, theoretical papers, case studies and book reviews. Special emphasis is given to articles that address models, diagnosis and remediation of reading, writing and spelling; orthography and its relation to reading and writing; computer literacy; cross cultural studies; and developmental and acquired disorders of reading and writing.

Subscription Information **ISSN** 0922-4777
1990, Volume 2 (4 issues)
Institutional rate: Dfl. 250.00/US$117.00
Incl. postage/handling
Private rate: Dfl. 106.00/US$50.00
Incl. postage/handling

Private subscriptions should be sent direct to the publishers.

P.O. Box 322, 3300 AH Dordrecht, The Netherlands
P.O. Box 358, Accord Station, Hingham, MA 02018-0358, U.S.A.

Journal Highlight

KLUWER ACADEMIC PUBLISHERS

British Journal of Developmental Psychology

Contents of Volume 8, Part 3, September 1990

Social background, phonological awareness and children's reading — *I. S. Raz & P. Bryant*

Sibling relationships in childhood: Links with friendships and peer relationships — *C. Stocker & J. Dunn*

SPECIAL SECTION: The nature of everyday science: A brief introduction — guest edited by Giyoo Hatano

The nature of everyday science: A brief introduction — *G. Hatano*

A developmental analysis of everyday topology in unschooled straw weavers — *G. B. Saxe & M. Gearhart*

A situated schema of proportionality — *A. D. Schliemann & T. Nunes*

Everyday and curriculum-based physics concepts — *I. Levin, R. S. Siegler, S. Druyan & R. Gardosh*

Young children's use of knowledge in everyday biology — *K. Inagaki*

Commentaries by M. Cole; P. Harris; R. Driver

The **British Journal of Developmental Psychology** (ISSN 0261-510X) is published four times a year and edited by Professor George Butterworth (University of Stirling). The price of volume 8 (1990) is £61.00 (US$112.00). Members and Foreign Affiliates of the Society and individual members of APA and other selected societies are entitled to purchase journals at preferential rates.

Subscriptions and inquiries to
The British Psychological Society
The Distribution Centre, Blackhorse Road, Letchworth, Herts. SG6 1HN

British Journal of Psychology

Contents of Volume 81, Part 3, 1990

Sternberg's mixed model applied to intermediate linear syllogisms: A mismatch — *M. T. Maybery*

Mayday for Maybery: A reply to an invalid critique of the mixed model of linear-syllogistic reasoning — *R. J. Sternberg*

Why is it difficult to put names to faces? *G. Cohen*

Comparing the effects of alcohol and intelligence on text recall and recognition — *E. A. Maylor, P. M. A. Rabbitt, G. H. James & S. A. Kerr*

The psychometric properties of the Bortner Type A Scale — *J. R. Edwards, A. J. Baglioni Jr & C. L. Cooper*

Name recall performance across the adult life-span — *T. H. Crook & R. L. West*

Social facilitation and driver behaviour — *J. S. Baxter, A. S. R. Manstead, S. G. Stradling, K. A. Campbell, J. T. Reason & D. Parker*

Understanding face recognition with an interactive activation model — *A. M. Burton, V. Bruce & R. A. Johnston*

The effects of photic stimulation and private self-consciousness on the complexity of visual imagination imagery — *A. Richardson & F. McAndrew*

The **British Journal of Psychology** (ISSN 0007-1269) is published four times a year and edited by Professor Antony J. Chapman (University of Leeds). The price of volume 81 (1990) is £97.00 (US$181.00).

Members and Foreign Affiliates of the Society and individual members of APA and other selected societies are entitled to purchase journals at preferential rates.

Subscriptions and inquiries to
The British Psychological Society
The Distribution Centre, Blackhorse Road, Letchworth, Herts. SG6 1HN

British Journal of Clinical Psychology

Contents of Volume 29, Part 3, September 1990

Behavioural and cognitive-behavioural treatments for epilepsy — *L. H. Goldstein*

Coping and help-seeking in the UK adult population — *C. Barker, N. Pistrang, D. A. Shapiro & I. Shaw*

Group counselling for relatives of hospitalized presenile dementia patients — *R. E. Perkins & C. F. Poynton*

Young adult offspring of parents with drinking problems: Recollections of parents' drinking and its immediate effects — *R. Velleman & J. Orford*

The relationship between the Eysenck Personality Questionnaire and Beck's concepts of sociotropy and autonomy — *P. Gilbert & S. Reynolds*

Eating disorders, perceived control, assertiveness and hostility — *G.-J. Williams, A. S. Chamove & H. R. Millar*

Factors underlying a psychiatric day hospital consumer survey — *N. A. Rothwell*

Clinical depression and attributional complexity — *G. L. Flett & P. L. Hewitt*

External attributions and outcome in depressive in-patients — *S. Priebe & R.-D. Stieglitz*

Psychological and psychophysiological characteristics in irritable bowel syndrome — *S. Thornton, P. McIntyre, I. Murray-Lyon & J. Gruzelier*

Dieting concerns of 10-year-old girls and their mothers — *A. J. Hill, C. Weaver & J. E. Blundell*

The **British Journal of Clinical Psychology** (ISSN 0144-6657) is published four times a year and edited by Professor Ray Hodgson (Whitchurch Hospital, Cardiff). The price of volume 29 (1990) is £65.00 (US$121.00).

Members and Foreign Affiliates of the Society and individual members of APA and other selected societies are entitled to purchase journals at preferential rates.

Subscriptions and inquiries to
The British Psychological Society
The Distribution Centre, Blackhorse Road, Letchworth, Herts. SG6 1HN, UK

British Journal of Medical Psychology

Contents of Volume 63, Part 3, September 1990

Obituary of Thomas Main

Concerning knowledge of mental states — *R. P. Hobson*

Interpreting mood scores: Clinical implications of individual differences in mood variability — *C. Cooper & C. McConville*

Women's first experiences of childbirth: Two hospital settings compared — *B. S. Bradley, M. Tashevska & J. M. Selby*

Transforming identities: A repertory grid case study of the transition to motherhood — *J. A. Smith*

Women's anticipation and experience of childbirth: The independence of fulfilment, unpleasantness and pain — *P. Salmon, R. Miller & N. C. Drew*

Maternal personality and sex of infant — *V. J. Grant*

Body image in rheumatoid arthritis: The relevance of hand appearance to desire for surgery — *M. Vamos*

The *in vivo* manipulation of alcohol-related beliefs in male social drinkers in a naturalistic setting — *T. P. S. Oei, J. Foley & R. McD. Young*

Authoritarianism as a cause of heart disease: Reply to Byrne, Reinhart & Heaven — *J. J. Ray*

The *British Journal of Medical Psychology* (ISSN 0007-1129) is published four times a year and edited by Dr John Birtchnell (MRC Social Psychiatry Unit, Institute of Psychiatry, London). The price of volume 63 (1990) is £76.00 (US$141.00).

Members and Foreign Affiliates of the Society and individual members of APA and other selected societies are entitled to purchase journals at preferential rates.

Subscriptions and inquiries to
The British Psychological Society
The Distribution Centre, Blackhorse Road, Letchworth, Herts. SG6 1HN

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

British Journal of Psychology
British Journal of Medical Psychology
Journal of Occupational Psychology
British Journal of Social Psychology
British Journal of Clinical Psychology
British Journal of Developmental Psychology
British Journal of Mathematical and Statistical Psychology
The British Journal of Educational Psychology

The *British Journal of Psychology*, the *British Journal of Medical Psychology*, *Journal of Occupational Psychology*, the *British Journal of Developmental Psychology*, the *British Journal of Social Psychology* and the *British Journal of Clinical Psychology* appear in quarterly parts. The 1990 subscription prices, payable in advance, are £97.00 (US$181.00) for the *British Journal of Psychology*; £65.00 (US$121.00) for the *British Journal of Social Psychology*; £65.00 (US$121.00) for the *British Journal of Clinical Psychology*; £61.00 (US$112.00) for the *British Journal of Developmental Psychology*; £76.00 (US$141.00) for the *British Journal of Medical Psychology* and £64.00 (US$119.00) for the *Journal of Occupational Psychology*.

The *British Journal of Mathematical and Statistical Psychology* is issued twice a year, in May and November. The subscription rate, payable in advance, is £64.00 per volume post free (US$119.00 overseas).

Subscriptions to any of these journals may be sent to a bookseller or subscription agent, or to The British Psychological Society, The Distribution Centre, Blackhorse Road, Letchworth, Herts., SG6 1HN, UK. A special price of £433.00 (US$805.00) for all seven journals going to the same address is offered for 1990.

The *British Journal of Educational Psychology* is issued three times a year, in February, June and November. The prepaid subscription is £30.00 a year post free, or £10.50 each number. The prepaid subscription price in North America is $60.00 US. Journals are sent accelerated surface post to the USA. Subscriptions should be sent to the Scottish Academic Press, 139 Leith Walk, Edinburgh, EH6 8NS.

MANUSCRIPTS

For the British Journal of Psychology should be sent to Professor A. J. Chapman, Journals Office, 13A Church Lane, London, N2 8DX.

For the British Journal of Medical Psychology should be sent to Dr John Birtchnell, Journals Office, 13A Church Lane, London N2 8DX.

For the Journal of Occupational Psychology should be sent to Dr Michael West, Journals Office, 13A Church Lane, London, N2 8DX

For the British Journal of Social Psychology should be sent to Dr Miles Hewstone, Journals Office, 13A Church Lane, London, N2 8DX.

For the British Journal of Clinical Psychology should be sent to Professor R. Hodgson, Journals Office, 13A Church Lane, London, N2 8DX

For the British Journal of Mathematical and Statistical Psychology should be sent to Dr A D. Lovie, Journals Office, 13A Church Lane, London, N2 8DX.

For the British Journal of Developmental Psychology should be sent to Professor George Butterworth, Journals Office, 13A Church Lane, London, N2 8DX.

For The British Journal of Educational Psychology should be sent in duplicate to Professor Charles Desforges, University of Exeter, St Luke's, Heavitree Road, Exeter, EX1 2LU.

Contributors are advised to consult the Style Guide issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price £3.50 post free, from The British Psychological Society, St Andrew's House, 48 Princes Road East, Leicester, LE1 7DR

All manuscripts should be submitted in duplicate.

OFFPRINTS

Contributors on application receive fifty copies of their papers free. Additional copies may be obtained in units of fifty copies.

REPRINTS

The Journal is available in reprint from Swets and Zeitlinger B.V., P.O. Box 810, 2160 SZ Lisse, The Netherlands; Vols. 21-29, 31-47.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £40.00.

They are entitled to purchase any or all of the Society's Journals at the special rates of £8.00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester, LE1 7DR.

BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY,

VOL. 60, PART 3, 1990

CONTENTS

Articles:

Säljö, R., and Wyndhamn, J. Problem-solving, Academic Performance and Situated Reasoning. A Study of Joint Cognitive Activity in the Formal Setting 245

Houghton, S., Wheldall, K., Jukes, R., and Sharpe, A. The Effects of Limited Private Reprimands and Increased Private Praise on Classroom Behaviour in Four British Secondary School Classes 255

Mason, J. The Development of Concepts about Written Language in the First Three Years of School 266

Sloper, P., Cunningham, C., Turner, S., and Knussen, C. Factors Related to the Academic Attainments of Children with Down's Syndrome 284

Van Overalle, F., and de Metsenaere, M. The Effects of Attribution-based Intervention and Study Strategy Training on Academic Achievement in College Freshmen 299

Norwich, B., and Duncan, J. Attitudes, Subjective Norm, Perceived Preventive Factors, Intentions and Learning Science: Testing a Modified Theory of Reasoned Action 312

King, N. J., Gullone, E., and Dadds, M.R. Student Perceptions of Permissiveness and Teacher-instigated Disciplinary Strategies 322

Plewis, I., Mooney, A., and Creeser, R. Time on Educational Activities at Home and Educational Progress in Infant School 330

Research Notes:

Wallon, P., and Baudoin, C. A New Criterium in the Assessment of the Behaviour of the Drawing Child 338

Leach, D. J., and Siddall, S.W. Parental Involvement in the Teaching of Reading: a Comparison of Hearing Reading, Paired Reading, Pause Prompt, Praise, and Direct Instruction Methods 349

Kember, D., and Gow, L. Cultural Specificity of Approaches to Study 356

Book Reviews 364

Letter to the Editor: The Burt Affair 375

Publications Received 376

Author Index for Volume 60 378

Editorial Acknowledgment 380

Notes to Contributors 381

UK ISSN 0007-0998